



CC and antagonists of *E. faecalis* protein activity  
XX  
SQ Sequence 627 AA;

Query Match 100.0%; Score 3225; DB 2; Length 627;  
Best Local Similarity 100.0%; Pred. No. 7.9e-198;  
Matches 627; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MQOLKKWYTVSTLLILPLFTSVLGTTTAPFAENGESQVLIHKKKMTDLDPPLIQNSG 60
DB 1 MQOLKKWYTVSTLLILPLFTSVLGTTTAPFAENGESQVLIHKKKMTDLDPPLIQNSG 60
QY 61 KEMSEFDKXQGLADYVFSIYNTNEFYEBORAAGASVDAKQAVOSLTPKQPVAAQGTDDAN 120
DB 61 KEMSEFDKXQGLADYVFSIYNTNEFYEBORAAGASVDAKQAVOSLTPKQPVAAQGTDDAN 120
QY 121 GNVTVQLPKKQNGKDAVYTIKSEPKKGVVAATNMVVAFPVYEMIKQTDGSKYKTEELAV 180
DB 121 GNVTVQLPKKQNGKDAVYTIKSEPKKGVVAATNMVVAFPVYEMIKQTDGSKYKTEELAV 180
QY 181 VHIYPRNVAVANDGSLHVKVGTAEENGNGAFVLSKSGSPGVYKIQGVNDGLYTWTT 240
DB 181 VHIYPRNVAVANDGSLHVKVGTAEENGNGAFVLSKSGSPGVYKIQGVNDGLYTWTT 240
QY 241 DKEQAKRPITGKSYELGENDPTEAENGTEGELTVKQLEVGSYLIEEVKADNNMELIENQTK 300
DB 241 DKEQAKRPITGKSYELGENDPTEAENGTEGELTVKQLEVGSYLIEEVKADNNMELIENQTK 300
QY 301 TPTTIEANNQTVVEKTVKNDTSKVDKTTPTSLDGKVAIGEKIKYQISVNIPLGIADKEGD 360
DB 301 TPTTIEANNQTVVEKTVKNDTSKVDKTTPTSLDGKVAIGEKIKYQISVNIPLGIADKEGD 360
QY 361 ANKYVKNLVNDGDAALTFDNTSGEYAYALVDGDTVIAPENYQVTEQANGFTVANPAY 420
DB 361 ANKYVKNLVNDGDAALTFDNTSGEYAYALVDGDTVIAPENYQVTEQANGFTVANPAY 420
QY 421 IPTLPFGTLKFVYFMHLINEKADPTKGFKNKANVNDGHTDDQTPPVEVVTGGKRIKXD 480
DB 421 IPTLPFGTLKFVYFMHLINEKADPTKGFKNKANVNDGHTDDQTPPVEVVTGGKRIKXD 480
QY 481 GNVTAQALAGASPVVRDQNSDTANYLKIDETTFKATWYKTAEAATFTTTADGLVDITG 540
DB 481 GNVTAQALAGASPVVRDQNSDTANYLKIDETTFKATWYKTAEAATFTTTADGLVDITG 540
QY 541 LKGYTYLLEETVAPDDYVLLTNRIEFVNVESYGTTEINVSPEKVPNKKHGLPSTGGKG 600
DB 541 LKGYTYLLEETVAPDDYVLLTNRIEFVNVESYGTTEINVSPEKVPNKKHGLPSTGGKG 600
QY 601 IYVYLGSGAVLLLIAGVYFARRRKENA 627
DB 601 IYVYLGSGAVLLLIAGVYFARRRKENA 627

RESULT 2
ABP43337
ID ABP43337 standard; protein; 627 AA.
XX ABP43337;
AC
XX
XX 05-AUG-2002 (first entry)
XX
XX E faecalis Bf058 protein.
DE
XX
XX Enterococcus; vaccine; gastrointestinal disease; diagnosis; antibiotic.
OS
XX Enterococcus faecalis.
XX
XX US2002045737-A1.
XX
XX 18-APR-2002.
XX
XX 04-MAY-1998; 98US-00071035.
XX

```

```

PR 04-MAY-1998; 98US-00071035.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Choi GH, Bailey C, Hromockyj A, Kunesh CA, G4480413
XX WPI; 2002-425450/45.
XX N-PSDB; ABN98093.
XX
XX New genes and polypeptides from Enterococcus faecalis, useful as vaccines
XX for preventing, treating or attenuating an infection caused by a member
XX of the Enterococcus genus in an animal, particularly E. faecalis.
XX
XX Claim 9; Page 93; 255bp; English.
XX
XX The present invention provides the protein and coding sequences of a
XX number of polypeptides from Enterococcus faecalis. The proteins can be
XX used as vaccines for preventing or attenuating an infection caused by a
XX member of the Enterococcus genus in an animal, particularly E. faecalis.
XX The polynucleotide is also useful for preventing or treating E. faecalis
XX infection. The present sequence is a protein of the invention
XX
XX Sequence 627 AA:
SQ
Query Match 100.0%; Score 3225; DB 5; Length 627;
Best Local Similarity 100.0%; Pred. No. 7.9e-198;
Matches 627; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 MQOLKKWYTVSTLLILPLFTSVLGTTTAPFAENGESQVLIHKKKMTDLDPPLIQNSG 60
DB 1 MQOLKKWYTVSTLLILPLFTSVLGTTTAPFAENGESQVLIHKKKMTDLDPPLIQNSG 60
QY 61 KEMSEFDKXQGLADYVFSIYNTNEFYEBORAAGASVDAKQAVOSLTPKQPVAAQGTDDAN 120
DB 61 KEMSEFDKXQGLADYVFSIYNTNEFYEBORAAGASVDAKQAVOSLTPKQPVAAQGTDDAN 120
QY 121 GNVTVQLPKKQNGKDAVYTIKSEPKKGVVAATNMVVAFPVYEMIKQTDGSKYKTEELAV 180
DB 121 GNVTVQLPKKQNGKDAVYTIKSEPKKGVVAATNMVVAFPVYEMIKQTDGSKYKTEELAV 180
QY 181 VHIYPRNVAVANDGSLHVKVGTAEENGNGAFVLSKSGSPGVYKIQGVNDGLYTWTT 240
DB 181 VHIYPRNVAVANDGSLHVKVGTAEENGNGAFVLSKSGSPGVYKIQGVNDGLYTWTT 240
QY 241 DKEQAKRPITGKSYELGENDPTEAENGTEGELTVKQLEVGSYLIEEVKADNNMELIENQTK 300
DB 241 DKEQAKRPITGKSYELGENDPTEAENGTEGELTVKQLEVGSYLIEEVKADNNMELIENQTK 300
QY 301 TPTTIEANNQTVVEKTVKNDTSKVDKTTPTSLDGKVAIGEKIKYQISVNIPLGIADKEGD 360
DB 301 TPTTIEANNQTVVEKTVKNDTSKVDKTTPTSLDGKVAIGEKIKYQISVNIPLGIADKEGD 360
QY 361 ANKYVKNLVNDGDAALTFDNTSGEYAYALVDGDTVIAPENYQVTEQANGFTVANPAY 420
DB 361 ANKYVKNLVNDGDAALTFDNTSGEYAYALVDGDTVIAPENYQVTEQANGFTVANPAY 420
QY 421 IPTLPFGTLKFVYFMHLINEKADPTKGFKNKANVNDGHTDDQTPPVEVVTGGKRIKXD 480
DB 421 IPTLPFGTLKFVYFMHLINEKADPTKGFKNKANVNDGHTDDQTPPVEVVTGGKRIKXD 480
QY 481 GNVTAQALAGASPVVRDQNSDTANYLKIDETTFKATWYKTAEAATFTTTADGLVDITG 540
DB 481 GNVTAQALAGASPVVRDQNSDTANYLKIDETTFKATWYKTAEAATFTTTADGLVDITG 540
QY 541 LKGYTYLLEETVAPDDYVLLTNRIEFVNVESYGTTEINVSPEKVPNKKHGLPSTGGKG 600
DB 541 LKGYTYLLEETVAPDDYVLLTNRIEFVNVESYGTTEINVSPEKVPNKKHGLPSTGGKG 600
QY 601 IYVYLGSGAVLLLIAGVYFARRRKENA 627
DB 601 IYVYLGSGAVLLLIAGVYFARRRKENA 627

```

RESULT 3  
ABU88365  
ID ABU88365 standard; protein; 627 AA.  
XX  
XX AC ABU88365;  
XX  
XX DT 07-JUL-2003 (first entry)  
XX  
XX DE E. faecalis novel protein #109.  
XX  
XX KW Vaccine; endocarditis; bacteraemia; urinary tract infection; UTI;  
XX intrabdominal infection; soft tissue infection; neonatal sepsis.  
XX  
XX OS Enterococcus faecalis.  
XX  
XX PN US2003017495-A1.  
XX  
XX PD 23-JAN-2003.  
XX  
XX PF 29-JUL-2002; 2002US-00206576.  
XX  
XX PR 06-MAY-1997; 97US-0044031P.  
XX 16-MAY-1997; 97US-004655P.  
XX 14-NOV-1997; 97US-0066009P.  
XX 04-MAY-1998; 98US-00071035.  
XX  
XX PA (HUMA-) HUMAN GENOME SCI INC.  
XX  
XX PI Choi GH, Bailey C, Hromockyj A, Kunsch CA;  
XX  
XX DR WPI: 2003-416890/39.  
XX  
XX DR N-PSDB: ACA88057.  
XX  
XX PT New nucleic acid molecules and polypeptides from Enterococcus faecalis,  
XX useful as vaccines for preventing or attenuating an enterococcal  
XX infection in an animal, or for identifying Enterococcus faecalis in  
XX biological samples.  
XX  
XX PS Claim 12; Page: 40pp; English.  
XX  
XX CC The invention relates to a new isolated nucleic acid molecule comprising  
XX a polynucleotide isolated from Enterococcus faecalis appearing as  
XX ACA87949-ACA88196 (or sequences complementary to them or 95% identical to  
XX them). Also included are the proteins encoded by the above nucleic acids,  
XX making a recombinant vector (comprising inserting the isolated nucleic  
XX acid molecule cited above into a vector), a host cell comprising the  
XX vector (used to produce the protein), an isolated antibody specific for  
XX the polypeptides, a hybridoma that produces the antibody, an isolated  
XX polypeptide antigen comprising an amino acid sequence of an Enterococcus  
XX faecalis epitope listed in the specification, a vaccine comprising one or  
XX more E. faecalis polypeptides (and a pharmaceutical diluent, carrier or  
XX excipient) where the polypeptide elicits protective antibodies in an  
XX animal to a member of the genus Enterococcus; preventing or attenuating  
XX an infection caused by a member of the genus Enterococcus in an animal  
XX comprising administering to the animal the polypeptide and detecting  
XX Enterococcus nucleic acids in a biological sample. The E. faecalis  
XX nucleic acid molecules and polypeptides are useful as vaccines for  
XX preventing or attenuating an enterococcal infection in an animal (e.g.  
XX endocarditis, bacteraemia, urinary tract infection (UTI), intrabdominal  
XX infection, soft tissue infection and neonatal sepsis). The polypeptides  
XX are also useful for detecting Enterococcus aureus in immunoassays, as  
XX epitope tags, as molecular weight markers, or for generating antibodies  
XX that specifically bind E. faecalis polypeptides. The nucleic acid  
XX molecules are also useful as probes for gene mapping, or for identifying  
XX E. faecalis in biological samples. The kit and methods are useful for  
XX detecting Enterococcus antibodies or nucleic acid molecules in a  
XX biological sample. The present sequence is a novel E. faecalis  
XX polypeptide of the invention. Note: The sequence data for this patent did  
XX not form part of the printed specification, but was obtained in  
XX electronic format directly from the USPTO at  
XX seqdata.uspto.gov/sequence.html?DocID=20030017495  
XX  
XX SQ Sequence 627 AA;

US Pat  
6448043

Query Match 100.0%; Score 3225; DB 6; Length 627;  
Best Local Similarity 100.0%; Pred. No. 7.9e-198;  
Matches 627; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MKQLKKWVYVSTLLILPLFTSVLGTTTAAEENGESAOYLHKKMDLDPDPLIONSG 60  
DB 1 MKQLKKWVYVSTLLILPLFTSVLGTTTAAEENGESAOYLHKKMDLDPDPLIONSG 60  
QY 61 KEMSEFDKYGGLADYTPSYNTNTEFEYRRAAGASVDAAKQVOSITPCKPAQCTTDAN 120  
DB 61 KEMSEFDKYGGLADYTPSYNTNTEFEYRRAAGASVDAAKQVOSITPCKPAQCTTDAN 120  
QY 121 GNVTVQLPKQNGKQAVVTIKKEPKEGVAAATNMVAFPVYEMIKQTDOSYKGTBEALAV 180  
DB 121 GNVTVQLPKQNGKQAVVTIKKEPKEGVAAATNMVAFPVYEMIKQTDOSYKGTBEALAV 180  
QY 181 VAIYPRNVVANDGSLHVKVGTAEENGNGAEFVLSKESGSPGVKYIQGVKQGLYTWTT 240  
DB 181 VAIYPRNVVANDGSLHVKVGTAEENGNGAEFVLSKESGSPGVKYIQGVKQGLYTWTT 240  
QY 241 DKEQAKRFTGKSYIEIGENDPTEAENGSELTVKULEVGSYILBEYKAPNNNELIENQTK 300  
DB 241 DKEQAKRFTGKSYIEIGENDPTEAENGSELTVKULEVGSYILBEYKAPNNNELIENQTK 300  
QY 301 TPFTTBANNQTPVEKTVKNDISKVDKTPSLDGKQVAIGEKIKYQISVNIPLGIADKEGD 360  
DB 301 TPFTTBANNQTPVEKTVKNDISKVDKTPSLDGKQVAIGEKIKYQISVNIPLGIADKEGD 360  
QY 361 ANKYVKFNLVDKGDALTPDNTSGEYAVALYDQGTVALPENYQVTEQANGFTVAVNPAY 420  
DB 361 ANKYVKFNLVDKGDALTPDNTSGEYAVALYDQGTVALPENYQVTEQANGFTVAVNPAY 420  
QY 421 ITTLTPGGTLKRVYFMHLEKADPTKGRNEANVNDGHTDDTPPVEYVYTGSKRPIKXD 480  
DB 421 ITTLTPGGTLKRVYFMHLEKADPTKGRNEANVNDGHTDDTPPVEYVYTGSKRPIKXD 480  
QY 481 GGVTAQNALAGSFVVRDONSPTANYLKIDETTRKATWTKAABATFTPTTADGLVDING 540  
DB 481 GGVTAQNALAGSFVVRDONSPTANYLKIDETTRKATWTKAABATFTPTTADGLVDING 540  
QY 541 LKGYTYLSEETVAPDDVYLLTNRIFVNVNBSYGTENLVSEKVPNKKHGLTPSTGKG 600  
DB 541 LKGYTYLSEETVAPDDVYLLTNRIFVNVNBSYGTENLVSEKVPNKKHGLTPSTGKG 600  
QY 601 IYVYLGSGAVLLIIGVTPARRRKENA 627  
DB 601 IYVYLGSGAVLLIIGVTPARRRKENA 627  
RESULT 4  
ABU13616  
ID ABU13616 standard; protein; 627 AA.  
XX  
XX AC ABU13616;  
XX  
XX DT 26-FEB-2003 (first entry)  
XX  
XX DE Enterococcus faecalis EF040 polypeptide #109.  
XX  
XX KW EF040; immunostimulant; antibacterial; gene mapping.  
XX  
XX OS Enterococcus faecalis.  
XX  
XX PN US6448043-B1.  
XX  
XX PD 10-SEP-2002.  
XX  
XX PF 04-MAY-1998; 98US-00071035.  
XX  
XX PR 06-MAY-1997; 97US-0044031P.  
XX 16-MAY-1997; 97US-004655P.  
XX 14-NOV-1997; 97US-0066009P.  
XX

```
PR 14-NOV-1997; 97US-0066099P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Choi GH, Bailey C, Hromockyj A, Kunsch CA;
XX
XX MPI; 2003-089120/08.
XX
XX N-PSDB; ABX61663.
XX
XX
XX New EF040 polypeptides and polynucleotides from Enterococcus faecalis,
XX useful for generating an immune response against E. faecalis and other
XX Enterococcus species, and as vaccines against other bacterial genera.
XX
XX Example 1; Col 115-116; 146pp; English.
XX
XX The invention relates to polynucleotide fragments of a gene from
XX Enterococcus faecalis, EF040, and the polypeptides encoded by them. The
XX polypeptides are useful in detecting E. faecalis, as epitope tags, as
XX molecular weight markers on SDS-PAGE gels or for molecular sieve gel
XX filtration columns, in generating antibodies that specifically bind to
XX the E. faecalis polypeptides, in generating an immune response against E.
XX faecalis and other Enterococcus species and as vaccines against other
XX bacterial genera. The polynucleotides are useful as probes for gene
XX mapping and for identifying E. faecalis in biological samples. Sequences
XX ABU13508-ABU13755 represent EF040 polypeptides of the invention. Note:
XX The sequence data for this patent can also be obtained from USPTO at
XX seqdata.uspto.gov/sequence.html
XX
XX Sequence 627 AA;
SQ
Query Match 100.0%; Score 3225; DB 6; Length 627;
Best Local Similarity 100.0%; Pred. No. 7.9e-198;
Matches 627; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MKQLKKWYTVSTLLILPLFTSVLGTTTAPAEENGESAQVLIHKKKMTDLPDPLIIONSG 60
DB 1 MKQLKKWYTVSTLLILPLFTSVLGTTTAPAEENGESAQVLIHKKKMTDLPDPLIIONSG 60
QY 61 KEMSEFDKXQGLADVTFSIYVNTNEFYBQRAAGASVDAKQAVQSLTPGKPAQGTDDAN 120
DB 61 KEMSEFDKXQGLADVTFSIYVNTNEFYBQRAAGASVDAKQAVQSLTPGKPAQGTDDAN 120
QY 121 GNVTVQLPRKQNGKDAVYTIKEEPKEGVAAATNMVVAFPVYEMIKQDGSYKYGTEELAV 180
DB 121 GNVTVQLPRKQNGKDAVYTIKEEPKEGVAAATNMVVAFPVYEMIKQDGSYKYGTEELAV 180
QY 181 VHIYKRVNVANDGSLHVKKVGTAEENGGLNGAEFVLSKSESGPGTVYITGVXDGLYTWT 240
DB 181 VHIYKRVNVANDGSLHVKKVGTAEENGGLNGAEFVLSKSESGPGTVYITGVXDGLYTWT 240
QY 241 DKEQAKRFTGKSYEIGENDPTEAENGTELTVKNLEVGSYLIEFKAPNNMELLENQTK 300
DB 241 DKEQAKRFTGKSYEIGENDPTEAENGTELTVKNLEVGSYLIEFKAPNNMELLENQTK 300
QY 301 TEPFTEANNQTPVEKTVKNDTSKVDKTPSLDQKQVAIGEIKRYQISVNIPLGIADKED 360
DB 301 TEPFTEANNQTPVEKTVKNDTSKVDKTPSLDQKQVAIGEIKRYQISVNIPLGIADKED 360
QY 361 ANKYKRVNLVNDKDAALFTDNVTSGRYAYALYDGTVAIPENYQTEQANGTVAVNPAY 420
DB 361 ANKYKRVNLVNDKDAALFTDNVTSGRYAYALYDGTVAIPENYQTEQANGTVAVNPAY 420
QY 421 IPTLPFGGTLKFVYFMHLENEKADPTKGFNGEANVDGHTDDQTPPVEVVTGSKAFIKYD 480
DB 421 IPTLPFGGTLKFVYFMHLENEKADPTKGFNGEANVDGHTDDQTPPVEVVTGSKAFIKYD 480
QY 481 GDNVTAQALAGASFYVRDQNSDPTANYLKIDETTKATVWTKRAEATTTTADGLVDING 540
DB 481 GDNVTAQALAGASFYVRDQNSDPTANYLKIDETTKATVWTKRAEATTTTADGLVDING 540
QY 541 LKYGTYVLEETVAPDDYVLLINRIEFVNVESYGTENIVSPBKIPNKKKGLPSTGKG 600
DB 541 LKYGTYVLEETVAPDDYVLLINRIEFVNVESYGTENIVSPBKIPNKKKGLPSTGKG 600
```

```
QY 601 IYVYIGSGAVLLILAGVYFARRRKENA 627
DB 601 IYVYIGSGAVLLILAGVYFARRRKENA 627
RESULT 5
ID ADY38952 standard; protein; 627 AA.
XX
XX ADY38952;
XX
XX 05-MAY-2005 (first entry)
XX
XX DE Novel Enterococcus faecalis protein sequence SeqID218.
XX
XX protein purification; DNA purification; antibacterial; vaccine;
XX enterococcus faecalis infection.
XX
XX Enterococcus faecalis.
XX
XX US2005043528-A1.
XX
XX 24-FEB-2005.
XX
XX 06-AUG-2004; 2004US-00912362.
XX
XX 06-MAY-1997; 97US-0044031P.
XX
XX 16-MAY-1997; 97US-0046555P.
XX
XX 14-NOV-1997; 97US-0066099P.
XX
XX 04-MAY-1998; 98US-00071035.
XX
XX 29-JUL-2002; 2002US-00206576.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Choi GH, Bailey C, Hromockyj A, Kunsch CA;
XX
XX MPI; 2005-180840/19.
XX
XX N-PSDB; ADY38951.
XX
XX PT New isolated Enterococcus faecalis nucleic acid useful for diagnosing,
XX preventing and/or attenuating infection by Enterococcus faecalis in
XX animals, particularly in humans.
XX
XX PS Claim 9; SEQ ID NO 218; 28pp; English.
XX
XX
XX This invention relates to novel genes and the proteins they encode
XX isolated from Enterococcus faecalis. The invention may be useful for the
XX development of compounds with an antibacterial activity or a vaccine. The
XX present invention is useful for the diagnosis, prevention and/or
XX attenuation of infection by Enterococcus faecalis in animals,
XX particularly in humans. The present sequence is that of a novel
XX Enterococcus faecalis protein of the invention. Note: The present
XX sequence does not form part of the printed specification but was obtained
XX in electronic form from the USPTO web.
XX
XX Sequence 627 AA;
SQ
Query Match 100.0%; Score 3225; DB 9; Length 627;
Best Local Similarity 100.0%; Pred. No. 7.9e-198;
Matches 627; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MKQLKKWYTVSTLLILPLFTSVLGTTTAPAEENGESAQVLIHKKKMTDLPDPLIIONSG 60
DB 1 MKQLKKWYTVSTLLILPLFTSVLGTTTAPAEENGESAQVLIHKKKMTDLPDPLIIONSG 60
QY 61 KEMSEFDKXQGLADVTFSIYVNTNEFYBQRAAGASVDAKQAVQSLTPGKPAQGTDDAN 120
DB 61 KEMSEFDKXQGLADVTFSIYVNTNEFYBQRAAGASVDAKQAVQSLTPGKPAQGTDDAN 120
QY 121 GNVTVQLPRKQNGKDAVYTIKEEPKEGVAAATNMVVAFPVYEMIKQDGSYKYGTEELAV 180
DB 121 GNVTVQLPRKQNGKDAVYTIKEEPKEGVAAATNMVVAFPVYEMIKQDGSYKYGTEELAV 180
```



QY 181 VHIYKGVVANDGSLHVKVKGTAENGLNGAEFVISKSESGPGTVKYGIVGKGLTYMTT 240  
 DB 181 VHIYKGVVANDGSLHVKVKGTAENGLNGAEFVISKSESGPGTVKYGIVGKGLTYMTT 240  
 QY 241 DKEQAKRPTGSKSYEIGENDPTEAENGTEGELTVKNLEVGSYLIEVKAPNNMELIENQTK 300  
 DB 241 DKEQAKRPTGSKSYEIGENDPTEAENGTEGELTVKNLEVGSYLIEVKAPNNMELIENQTK 300  
 QY 301 TPEFTIANNQTPVEKTVKNDTSKVDTKTTSLDCKDVAIGBKIKYQISVNIPLGIADKEGD 360  
 DB 301 TPEFTIANNQTPVEKTVKNDTSKVDTKTTSLDCKDVAIGBKIKYQISVNIPLGIADKEGD 360  
 QY 361 ANKYVFNVLVDKDALTFDNTVSGEYAVALYDGDVIAIPENYQVTEQANGFTVAVNPAY 420  
 DB 361 ANKYVFNVLVDKDALTFDNTVSGEYAVALYDGDVIAIPENYQVTEQANGFTVAVNPAY 420  
 QY 421 IPTLTTPGGTLKRVYFMHLINEKADPTKGFKNBANVDNGHTDDQTPPVEVVTGKRFIKVD 480  
 DB 421 IPTLTTPGGTLKRVYFMHLINEKADPTKGFKNBANVDNGHTDDQTPPVEVVTGKRFIKVD 480  
 QY 481 GGVTAIOALAGASFVVRDONSPTANYLKI DETTKAATWVKTAKAATTTTADGLVDITG 540  
 DB 481 GGVTAIOALAGASFVVRDONSPTANYLKI DETTKAATWVKTAKAATTTTADGLVDITG 540  
 QY 541 LKYGTYLLEETVAPDDVYLLTNRIEFVNVNOSYGTTENLVSPKVPNKHGTLPTSGGK 600  
 DB 541 LKYGTYLLEETVAPDDVYLLTNRIEFVNVNOSYGTTENLVSPKVPNKHGTLPTSGGK 600  
 QY 601 IYVYLGSAVLLLLIAGVYFARRRKENA 627  
 DB 601 IYVYLGSAVLLLLIAGVYFARRRKENA 627

## RESULT 6

ADH8239 standard; protein; 659 AA.

ADH8239;

22-APR-2004 (first entry)

Enterococcus faecalis polypeptide #2719.

Enterococcus faecalis infection; transcription regulatory element;

antibacterial.

Enterococcus faecalis.

US6617156-B1.

09-SEP-2003

13-AUG-1996; 98US-00134000.

15-AUG-1997; 97US-0055778P.

(DOUC/) DOUCETTE-STAMM L A.

(BUSH/) BUSH D.

Doucette-Stamm LA, Bush D;

WPI; 2003-895394/82.

N-PSDB; ADH84834.

New nucleic acid comprising a sequence encoding an Enterococcus faecalis

polypeptide, useful for preparing a composition for diagnosing or

treating E. faecalis infection.

Disclosure; SEQ ID NO 6124; 193pp; English.

The invention relates to Enterococcus faecalis polynucleotides and

CC vector comprising a polynucleotide operably linked to a transcription  
 CC regulatory element, a cell comprising a recombinant vector, a method for  
 CC producing an E. faecalis polypeptide, an isolated nucleic acid comprising  
 CC a sequence not given in the specification, a recombinant vector  
 CC comprising the nucleic acid and a cell comprising the recombinant vector.  
 CC The polynucleotides can be used to detect the presence of E. faecalis in  
 CC a sample. The sequences are useful for preparing a composition for  
 CC diagnosing or treating Enterococcus faecalis infection. This sequence  
 CC represents an E. faecalis polypeptide of the invention.

SQ Sequence 659 AA;

Query Match 100.0%; Score 3225; DB 7; Length 659;

Best Local Similarity 100.0%; Pred. No. 8,5e-198;

Matches 627; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKQLKVVWTVSTLLILPLFTSVLGTTTAFABENGESQOLVHKKMTDLPDPLIQNSG 60  
 DB 33 MKQLKVVWTVSTLLILPLFTSVLGTTTAFABENGESQOLVHKKMTDLPDPLIQNSG 92  
 QY 61 KEMSEPDKYQGLADVTPSIVYNTNPEYEBORAGASVDAKQAVOSLTPKQVPAQGTIDAN 120  
 DB 93 KEMSEPDKYQGLADVTPSIVYNTNPEYEBORAGASVDAKQAVOSLTPKQVPAQGTIDAN 152  
 QY 121 GNVTVQLPKQNGKDAVYTIKEEPKGVAAATNMVAFVYEMIKQTDGSKYKTEELAV 180  
 DB 153 GNVTVQLPKQNGKDAVYTIKEEPKGVAAATNMVAFVYEMIKQTDGSKYKTEELAV 212  
 QY 181 VHIYKGVVANDGSLHVKVKGTAENGLNGAEFVISKSESGPGTVKYGIVGKGLTYMTT 240  
 DB 213 VHIYKGVVANDGSLHVKVKGTAENGLNGAEFVISKSESGPGTVKYGIVGKGLTYMTT 272  
 QY 241 DKEQAKRPTGSKSYEIGENDPTEAENGTEGELTVKNLEVGSYLIEVKAPNNMELIENQTK 300  
 DB 273 DKEQAKRPTGSKSYEIGENDPTEAENGTEGELTVKNLEVGSYLIEVKAPNNMELIENQTK 332  
 QY 301 TPEFTIANNQTPVEKTVKNDTSKVDTKTTSLDCKDVAIGBKIKYQISVNIPLGIADKEGD 360  
 DB 333 TPEFTIANNQTPVEKTVKNDTSKVDTKTTSLDCKDVAIGBKIKYQISVNIPLGIADKEGD 392  
 QY 361 ANKYVFNVLVDKDALTFDNTVSGEYAVALYDGDVIAIPENYQVTEQANGFTVAVNPAY 420  
 DB 393 ANKYVFNVLVDKDALTFDNTVSGEYAVALYDGDVIAIPENYQVTEQANGFTVAVNPAY 452  
 QY 421 IPTLTTPGGTLKRVYFMHLINEKADPTKGFKNBANVDNGHTDDQTPPVEVVTGKRFIKVD 480  
 DB 453 IPTLTTPGGTLKRVYFMHLINEKADPTKGFKNBANVDNGHTDDQTPPVEVVTGKRFIKVD 512  
 QY 481 GGVTAIOALAGASFVVRDONSPTANYLKI DETTKAATWVKTAKAATTTTADGLVDITG 540  
 DB 513 GGVTAIOALAGASFVVRDONSPTANYLKI DETTKAATWVKTAKAATTTTADGLVDITG 572  
 QY 541 LKYGTYLLEETVAPDDVYLLTNRIEFVNVNOSYGTTENLVSPKVPNKHGTLPTSGGK 600  
 DB 573 LKYGTYLLEETVAPDDVYLLTNRIEFVNVNOSYGTTENLVSPKVPNKHGTLPTSGGK 632  
 QY 601 IYVYLGSAVLLLLIAGVYFARRRKENA 627  
 DB 633 IYVYLGSAVLLLLIAGVYFARRRKENA 659

## RESULT 7

ADH84889 standard; protein; 625 AA.

ADH84889;

29-JUL-2004 (first entry)

E faecalis surface anchored LpxTG protein Seqid13.

LpxTG; cell wall-anchored surface protein; Gram positive bacterium;

extracellular matrix molecule; sequence database; C-terminal.

KM Immunoglobulin-like fold region; Ig-like fold region; antibacterial;  
 KM vaccine; gene therapy; infection; medical device; prosthesis;  
 KM premature newborn; AIDS; debilitated cancer; bone marrow transplantation.  
 XX  
 OS *Enterococcus faecalis*.  
 XY

~~WO200402546-A2.~~

25-MAR-2004

15-SEP-2003; 2003WO-US028789.

13-SEP-2002; 2002US-0410303P.

(TEXA) UNIV TEXAS A & M SYSTEM.

(UABR-) UAB RES FOUND.

Hook M, Xu Y, Sillanpää J.

WDT 0004 31EC04/00

N-PSDB; ADO84888.

PT Identifying LPTXG-containing cell wall-anchored surface proteins from  
PT Gram positive bacteria, for treating infection caused by the bacteria,  
PT PT comprises searching sequence information database for the sequence having  
PT LPTXG-motif.

PS Claim 16; SEQ ID NO 13; 96pp; English.

This invention relates to a novel method of identifying LPXTG-containing cell wall-anchored surface proteins from Gram positive bacteria that bind to an extracellular matrix molecule which comprises searching a database of sequence information for a putative protein sequence having the LPXTG motif in its C-terminal region and analysing the sequence for the presence of one or more Immunoglobulin (Ig)-like fold regions. The invention may be useful for the production of compounds with an antibacterial activity or for production of a vaccine. In addition the disclosed sequences may be useful for gene therapy. The antibody is useful for treating or preventing an infection of Gram-positive bacteria in a human or animal patient. The method and the proteins are useful in generating antibodies for treating and preventing the spread of infections of Gram positive bacteria, for interfering with, or inhibiting binding interactions by Gram positive bacteria, for monitoring the level of gram positive bacterial antigens, or antibodies recognising the antigens in a human or animal patients suspected of containing the antigens or antibodies, in preventing or reducing infection of medical devices and prostheses caused by such organisms, and in treating or preventing infections in highly susceptible groups such as premature newborns, AIDS and debilitated cancer patients, and bone marrow transplantation. The present sequence is that of a surface anchored LPXTG protein identified using the method of the invention.

SQ Sequence 625 AA;

|             |       |             |       |             |
|-------------|-------|-------------|-------|-------------|
| Query Match | 99.3% | Score 3202; | DB 8; | Length 625; |
|-------------|-------|-------------|-------|-------------|

Matches 625; Conservative 0; Mismatches 0; Indels 2; Gaps 1

QY 1 MKQKKWYTVSTLLILPLFTSVLGTTTAFABENGESAQLVHKKKMTDLPDPLIÖNSG 60

|    |     |                                                               |     |
|----|-----|---------------------------------------------------------------|-----|
| Db | 1   | KQQLKKWWTYSTLLILFLFTSVLGTTTAAFAENENESQVLVHKKKMDLPPRLIÖNSG     | 60  |
| Qy | 61  | KEMSEFDKXQGLADYVFSIYNTNNEFYEQRAGASVDAAKQVSLTPGRPAQGTDDAN      | 120 |
| Db | 61  | KEMSEFDKXQGLADYVFSIYNTNNEFYEQRAAGASVDAAKQVSLTPGRPAQGTDDAN     | 120 |
| Qy | 121 | GNVTVQLPKKÖNGKDAVYTIKEEPKEGVAAATNMVVAFPVYEMIÖÖTDSYKKGTEELAV   | 180 |
| Db | 121 | GNVTVQLPKKÖNGKDAVYTIKEEPKEGVAAATNMVVAFPVYEMIÖÖTDSYKKGTEELAV   | 180 |
| Qy | 181 | VHIYPRKNVANDOSLHKVKGTAENEGNGLANGAEFVISKSGSPGTAKYIÖGVKQGLYTWTT | 240 |

|    |     |                                                              |     |
|----|-----|--------------------------------------------------------------|-----|
| Db | 181 | VHILPYGNVANDGSL--KKVGTAEINEGNLGAEPVLSKSEGSPGVTKYIQGVKGLYTWTT | 238 |
| Qy | 241 | DKEQARKFTTGSYSYELGENDPTEANNGELTVKULEVGSYILEEYKAPNNALIEINQTK  | 300 |
| Db | 239 | DKEQARKFTTGSYSYELGENDPTEANNGELTVKULEVGSYILEEYKAPNNALIEINQTK  | 298 |
| Qy | 301 | TPETIEANNQTPPEKTVKNDTSKVQKTPSLDGDQVALGEKIKYQISVNIPLGIADKEGD  | 366 |
| Db | 299 | TPETIEANNQTPPEKTVKNDTSKVQKTPSLDGDQVALGEKIKYQISVNIPLGIADKEGD  | 358 |
| Qy | 361 | ANKYKFNVLVDKHDALTFEDNTSGEYAYALYDGDVTVAPENNQVTEQANGFTVAANPAY  | 420 |
| Db | 359 | ANKYKFNVLVDKHDALTFEDNTSGEYAYALYDGDVTVAPENNQVTEQANGFTVAANPAY  | 418 |
| Qy | 421 | IPTLTGGLTKFVYFMHLINEKADPTKGFKNZANVNGHTDDQTPPYEVVYTGGRPIKVD   | 480 |
| Db | 419 | IPTLTGGLTKFVYFMHLINEKADPTKGFKNZANVNGHTDDQTPPYEVVYTGGRPIKVD   | 478 |
| Qy | 481 | GDVTAFOALAGASFVVRDQNSDTANYLKIDETTKAATVTKKAATFTTTAAGLVNDITG   | 540 |
| Db | 479 | GDVTAFOALAGASFVVRDQNSDTANYLKIDETTKAATVTKKAATFTTTAAGLVNDITG   | 538 |
| Qy | 541 | LKKGTYLLEETVAPDDYVLTNRIRIEPVNNEOSYGTENLVSPKAVPKHGKTLPTSTGKG  | 600 |
| Db | 539 | LKKGTYLLEETVAPDDYVLTNRIRIEPVNNEOSYGTENLVSPKAVPKHGKTLPTSTGKG  | 598 |
| Qy | 601 | IYVYLSGAVLLIAGVYFARRRKKENA627                                |     |
| Db | 599 | IYVYLSGAVLLIAGVYFARRRKKENA625                                |     |

RESULT: 8  
ADY16503

|    |          |                           |
|----|----------|---------------------------|
| ID | ADV16503 | standard; protein; 625 AA |
|----|----------|---------------------------|

AC ADV165037

DT 24-FEB-2005 (first entry)

DE E. faecalis V583 hyperimmune serum reactive antigen protein - SEQ ID 198.  
XX  
XX antigen; antibacterial; vaccine; enterococcus infection; infection;  
KW pharyngitis; impetigo; rheumatic fever; antipyretic; antirheumatic;  
KM immunosuppressive; sepsis.

OS Enterococcus faecalis V583.

| FH | Key | Location/Qualifiers |
|----|-----|---------------------|
|    |     |                     |

FT /note= "Serum reactive epitope; claimed"

```
FT      /note= "Serum reactive epitope; claimed"
```

```
FT      /note= "Serum reactive epitope; claimed"
```

```

FT      /note= "Serum reactive epitope; claimed

```

```
FT /note= "Serum reactive epitope; claimed"
E
```

```

FT      /note= "Serum reactive epitope; claimed
      00  103

```

| FT  | note= "Serum reactive epitope; claimed |
|-----|----------------------------------------|
| 104 | 100                                    |

```
FT      /note= "Serum reactive epitope; claimed
100      100
```

| FT  | note= "Serum reactive epitope; claimed |
|-----|----------------------------------------|
| 125 | 140                                    |

|    |     |     |                                         |
|----|-----|-----|-----------------------------------------|
| FT | 145 | 153 | /note= "Serum reactive epitope; claimed |
|----|-----|-----|-----------------------------------------|

| FT  | note= "Serum reactive epitope; claimed |
|-----|----------------------------------------|
| 177 | 103                                    |

FT /note= "Serum reactive epitope; claimed

FT Region 209..215  
/note="Serum reactive epitope; claimed"  
FT Region 221..232  
/note="Serum reactive epitope; claimed"  
FT Region 271..284  
/note="Serum reactive epitope; claimed"  
FT Region 331..337  
/note="Serum reactive epitope; claimed"  
FT Region 341..352  
/note="Serum reactive epitope; claimed"  
FT Region 360..378  
/note="Serum reactive epitope; claimed"  
FT Region 383..390  
/note="Serum reactive epitope; claimed"  
FT Region 392..401  
/note="Serum reactive epitope; claimed"  
FT Region 409..422  
/note="Serum reactive epitope; claimed"  
FT Region 428..435  
/note="Serum reactive epitope; claimed"  
FT Region 462..470  
/note="Serum reactive epitope; claimed"  
FT Region 474..480  
/note="Serum reactive epitope; claimed"  
FT Region 482..496  
/note="Serum reactive epitope; claimed"  
FT Region 531..538  
/note="Serum reactive epitope; claimed"  
FT Region 541..549  
/note="Serum reactive epitope; claimed"  
FT Region 551..560  
/note="Serum reactive epitope; claimed"  
FT Region 562..569  
/note="Serum reactive epitope; claimed"  
FT Region 576..582  
/note="Serum reactive epitope; claimed"  
FT Region 598..618  
/note="Serum reactive epitope; claimed"  
FT Region  
/note="Serum reactive epitope; claimed"  
FN WO2004106367-A2.  
XX  
XX 09-DEC-2004.  
XX  
XX 26-MAY-2004; 2004WO-EP005664.  
XX  
XX 30-MAY-2003; 2003EP-00450137.  
XX  
XX (INTE-) INTERCELL AG.  
XX  
XX  
XX Meinke A, Nagy E, Hanner M, Gelmann D;  
XX  
XX WPI, 2005-039707/04.  
XX  
XX N-PSDB; ADV16333.  
XX  
XX  
XX Novel isolated nucleic acid molecule encoding hyperimmune serum reactive  
XX antigen e.g., EP0020, EP0032, EP0062, EP0149, EP0196, EP0253, EP0270,  
XX EP0298 or its fragment, useful for producing vaccine against enterococcal  
XX infection.  
XX  
XX Example 3; SEQ ID NO 198; 175pp; English.  
XX  
XX  
XX The invention relates to a novel isolated nucleic acid molecule encoding  
XX a hyperimmune serum reactive antigen e.g. SEQ ID 171-340, 357-372 or 425-  
XX 476, or its fragment. The antigen of the invention demonstrates  
XX antibacterial activities and may be useful for producing a pharmaceutical  
XX preparation, particularly a vaccine, against enterococcal infection.  
XX Conditions associated with bacterial infection which may be prevented or  
XX treated include bacterial pharyngitis, scarlet fever, impetigo, rheumatic  
XX fever, necrotizing fasciitis and sepsis in humans. The current sequence  
XX is that of an Enterococcus faecalis V583 hyperimmune serum reactive  
XX antigen protein of the invention.  
XX  
XX Sequence 625 AA;  
SQ

Query Match 99.3%; Score 3202; DB 9; Length 625;  
Best Local Similarity 99.7%; Pred. No. 2,3e-196;  
Matches 625; Conservative 0; Mismatches 0; Indels 2; Gaps 1;  
QY 1 MKQLKKWYVSTLLILPLFTSVIGTTTAPAEENGSAQVIAHKKMTDLPDPLIQNSG 60  
DB 1 MKQLKKWYVSTLLILPLFTSVIGTTTAPAEENGSAQVIAHKKMTDLPDPLIQNSG 60  
QY 61 KEMSEFDKXQGLADYTFPSIYNTNFEYBQRAAGASVDAKQVQSLTPKPAQGTDDN 120  
DB 61 KEMSEFDKXQGLADYTFPSIYNTNFEYBQRAAGASVDAKQVQSLTPKPAQGTDDN 120  
QY 121 GNVTVQLPKKQNGKQAVTIKEEPKGGVVAATNMVAFPMYMIKQTDSSYKGTBELAV 180  
DB 121 GNVTVQLPKKQNGKQAVTIKEEPKGGVVAATNMVAFPMYMIKQTDSSYKGTBELAV 180  
QY 181 VHIYPRNVVANDSLHVKKVGTAENEGNGAEFVLSKSESGDGVKYIQGVKGLYTWTT 240  
DB 181 VHIYPRNVVANDSLHVKKVGTAENEGNGAEFVLSKSESGDGVKYIQGVKGLYTWTT 240  
QY 241 DKEQAKRPITGKSYSIENGDFTEANGGELTVKULFQGSYILBKYKAPNNABELIENQK 300  
DB 241 DKEQAKRPITGKSYSIENGDFTEANGGELTVKULFQGSYILBKYKAPNNABELIENQK 300  
QY 301 TPFTTBANNQTPVEKTVKNDTSKVDKTPSLDGKQVAGKIKYQISVNIPLGIADKEGD 360  
DB 301 TPFTTBANNQTPVEKTVKNDTSKVDKTPSLDGKQVAGKIKYQISVNIPLGIADKEGD 360  
QY 361 ANKYVKNLVNDGHDALTFPDNVTSGEYVAYLYDGDVIAPENYQVTEQANGFTVAVNPAY 420  
DB 361 ANKYVKNLVNDGHDALTFPDNVTSGEYVAYLYDGDVIAPENYQVTEQANGFTVAVNPAY 420  
QY 421 IPTLPFGTLKRVYFMHLEKADPTKGRNEANVNDGHTDDTPTPVEVYVYTGKRPKID 480  
DB 421 IPTLPFGTLKRVYFMHLEKADPTKGRNEANVNDGHTDDTPTPVEVYVYTGKRPKID 480  
QY 439 IPTLPFGTLKRVYFMHLEKADPTKGRNEANVNDGHTDDTPTPVEVYVYTGKRPKID 478  
DB 439 IPTLPFGTLKRVYFMHLEKADPTKGRNEANVNDGHTDDTPTPVEVYVYTGKRPKID 478  
QY 481 GGVTVQALAGASFVVRDONSPTANYLKIDETTRKATWKTAKBATFTTTADGLVDING 540  
DB 479 GGVTVQALAGASFVVRDONSPTANYLKIDETTRKATWKTAKBATFTTTADGLVDING 538  
QY 541 LKGYTYLEETVAPDDVYLLTNRIEFVNEQSYGTENLVSPKVPNKKHGTLPSTGKG 600  
DB 539 LKGYTYLEETVAPDDVYLLTNRIEFVNEQSYGTENLVSPKVPNKKHGTLPSTGKG 598  
QY 601 IYVYLSGAVLLIAGVTPARRRKENA 627  
DB 599 IYVYLSGAVLLIAGVYFARRRKENA 625  
RESULT 9  
AAV00119  
ID AAV00119 standard; protein; 560 AA.  
XX  
XX AAV00119;  
AC  
XX  
XX 20-APR-1999 (first entry)  
DE Enterococcus faecalis antigenic polypeptide fragment EP058.  
XX  
XX Enterococcus faecalis; infection; vaccine; immune response; diagnosis;  
XX detection; attenuation; antigenic.  
OS Enterococcus faecalis.  
XX  
XX WO9850554-A2.  
PN  
XX  
XX 12-NOV-1998.  
PD  
XX  
XX 04-MAY-1998; 98WO-US008959.  
PF  
XX  
XX 06-MAY-1997; 97US-0044031P.  
PR  
XX  
XX 16-MAY-1997; 97US-0046555P.

|                       |                                                                           |                |
|-----------------------|---------------------------------------------------------------------------|----------------|
| DR                    | 14-NOV-1997;                                                              | 97US-0066009P. |
| XX                    | (HUMA-) HUMAN GENOME SCI INC.                                             |                |
| PA                    | Kunsch CA, Choi GH, Bailey C, Hromockyj A;                                |                |
| P1                    | WPI; 1999-070095/06.                                                      |                |
| DR                    | N-PSTDB; AAIX20109.                                                       |                |
| XX                    | New isolated Enterococcus faecalis polynucleotides - used to develop      |                |
| PT                    | products for the detection of Enterococcus and for use in vaccines for    |                |
| PT                    | prevention or attenuation of Enterococcus infection.                      |                |
| PS                    | Claim 9; Page 134; 301pp; English.                                        |                |
| XX                    | The present sequence represents an antigenic polypeptide fragment         |                |
| CC                    | isolated from Enterococcus faecalis. The present invention describes      |                |
| CC                    | genes, proteins and antigenic polypeptides isolated from E. faecalis. The |                |
| CC                    | proteins can be used in vaccines for preventing or attenuating an         |                |
| CC                    | infection caused by a member of the Enterococcus genus in an animal. They |                |
| CC                    | can also be used for detecting Enterococcus antibodies in a sample. The   |                |
| CC                    | nucleotide sequences can be used for detecting Enterococcus nucleic       |                |
| CC                    | acids. Products from the present invention can also be used for screening |                |
| CC                    | compounds to identify agonists and antagonists of E. faecalis protein     |                |
| CC                    | activity                                                                  |                |
| SQ                    | Sequence 560 AA;                                                          |                |
| Query Match           | 89.7%; Score 2892; DB 2; Length 560;                                      |                |
| Best Local Similarity | 100.0%; Pred. No. 1.4e-176;                                               |                |
| Matches               | 560; Conservative 0; Mismatches 0; Indels 0; Gaps 0                       |                |
| OY                    | 33 EENGSSAOLVIHKKKOTDLPDLIUNGSGEMSEFPDKYOGADVTFSIYNVTNEFFEORAA            | 92             |
| Db                    | 1 EENGSSAQVLHIKKKOTDLPDLIUNGSGEMSEFPDKYOGADVTFSIYNVTNEFFEORAA             | 60             |
| OY                    | 93 GASVDAAQAQAVOSLTPGKPPVAQGTTDANGNTVQLPKONGDAVYTIKEEPKEGVAAAT            | 152            |
| Db                    | 61 GASVDAAQAQAVOSLTPGKPPVAQGTTDANGNTVQLPKONGDAVYTIKEEPKEGVAAAT            | 120            |
| OY                    | 153 NMVVAFFPYEEMI KOTDGSKYGTEELAAVHHIPKNVVANDGSLHYKKVGTAENEGNGAE          | 212            |
| Db                    | 121 NMVVAFFPYEEMI KOTDGSKYKITEELAAVHHIPKNVVANDGSLHYKKVGTAENEGNGAE         | 180            |
| OY                    | 213 FVISKSBSPGTVKYIQGVKGLYTWTDDKEOAKRFTTGSYEIGENDFTEANGTGELT              | 272            |
| Db                    | 181 FVISKSBSPGTVKYIQGVKGLYTWTDDKEOAKRFTTGSYEIGENDFTEANGTGELT              | 240            |
| OY                    | 273 VKNLEVGSIILIEEVKAPNNALIELNOTKPTFIETANNOPVEKTYKVNDSKYDKTTPSLD          | 332            |
| Db                    | 241 VKNLEVGSIILIEEVKAPNNALIELNOTKPTFIETANNOPVEKTYKVNDSKYDKTTPSLD          | 300            |
| OY                    | 333 GKDVAIIGEKIKYQISVINIPLGIADREGDANKYVKFNLDKHDAAALTDFDNVTSGEVAYALY       | 392            |
| Db                    | 301 GKDVAIIGEKIKYQISVINIPLGIADREGDANKYVKFNLDKHDAAALTDFDNVTSGEVAYALY       | 360            |
| OY                    | 393 DGDVTIAPBNTIOVTEOANGFTYAANPAYIPTLLPGGTLKFFVPMHLNEKADPTKGKFNEA         | 452            |
| Db                    | 361 DGDVTIAPBNTIOVTEOANGFTYAANPAYIPTLLPGGTLKFFVPMHLNEKADPTKGKFNEA         | 420            |
| OY                    | 453 NVNDNGHTDOQPPTYEVVTVGGKRFLKVDSDVATAOLAGSPVRNDSDTNLYKIDBT              | 512            |
| Db                    | 421 NVNDNGHTDOQPPTYEVVTVGGKRFLKVDSDVATAOLAGSPVRNDSDTNLYKIDBT              | 480            |
| OY                    | 513 TKAAITWKTKABATTFTTTADGLVIDITGLKYGTYYILEFTVAADDVYLTLNRIEPFVNEQS        | 572            |
| Db                    | 481 TKAAITWKTKABATTFTTTADGLVIDITGLKYGTYYILEFTVAADDVYLTLNRIEPFVNEQS        | 540            |
| OY                    | 573 YGTENTLVSPKVPNKHGKT 592                                               |                |
| Db                    | 541 YGTENTLVSPKVPNKHGKT 560                                               |                |

| Query                 | Match           | 89.7%                 | Score 2892;                                | DB 5;  | Length 560; |
|-----------------------|-----------------|-----------------------|--------------------------------------------|--------|-------------|
| Best Local Similarity | 100.0%          |                       | Pred. No. 1.4e-176;                        |        |             |
| Matches 560;          | Conservative 0; | Mismatches 0;         | Indels 0;                                  | Gaps 0 |             |
| DB                    | 1               | ENGGSSAQLVHKKKMTDLPDL | IONSGKEMSEFPKYGSLADVTSSINVTNPFYHORA        | 92     |             |
| QY                    | 33              | ENGSSAQLVHKKKMTDLPDL  | IONSGKEMSEFPKYGSLADVTSSINVTNPFYHORA        | 92     |             |
| DB                    | 1               | ENGGSSAQLVHKKKMTDLPDL | IONSGKEMSEFPKYGSLADVTSSINVTNPFYHORA        | 60     |             |
| QY                    | 93              | GASVDAARQAVQSLTPGKRV  | AGSTTDANGNVVOLPKQNGDAVYTIKEEPKEGVAA        | 152    |             |
| DB                    | 61              | GASVDAARQAVQSLTPGKRV  | AGSTTDANGNVVOLPKQNGDAVYTIKEEPKEGVAA        | 120    |             |
| QY                    | 153             | MMVAVAPPYEMTKQTDG     | SYKGTBELAVVHIYPKVVANDGSLHYKVGTAENEGINGAE   | 212    |             |
| DB                    | 121             | MMVAVAPPYEMTKQTDG     | SYKGTBELAVVHIYPKVVANDGSLHYKVGTAENEGINGAE   | 180    |             |
| QY                    | 213             | FVTSSEBSPGVVKYIOGV    | KDGLYMTTDKEQAKRFITGSKSYEIGENDPTEANGTGELT   | 272    |             |
| DB                    | 181             | FVTSSEBSPGVVKYIOGV    | KDGLYMTTDKEQAKRFITGSKSYEIGENDPTEANGTGELT   | 240    |             |
| QY                    | 273             | VKNLEVGSYLIEEVKAP     | NNAELIENQTKPTTIEANNQTPVEKTVKNDTSKYDKTTPSLD | 332    |             |
| DB                    | 241             | VKNLEVGSYLIEEVKAP     | NNAELIENQTKPTTIEANNQTPVEKTVKNDTSKYDKTTPSLD | 300    |             |
| QY                    | 333             | GQDVAVIGKIKYQISVNI    | PLGIADGEGDANKVKNLVDPKHAAITPDNVTSGEVAYALY   | 392    |             |
| DB                    | 301             | GQDVAVIGKIKYQISVNI    | PLGIADGEGDANKVKNLVDPKHAAITPDNVTSGEVAYALY   | 360    |             |
| QY                    | 393             | DGDVTAIEENTQVTSQANG   | FVAVNPAYIPTLTGGTLKTVVFMHINERAKDPTKGFQKNEA  | 452    |             |



```

XX OS Enterococcus faecalis.
XX PN US6448043-B1.
XX PD 10-SEP-2002.
XX PF 04-MAY-1998; 98US-00071035.
XX PR 06-MAY-1997; 97US-0044031P.
XX PR 16-MAY-1997; 97US-004655P.
XX PR 14-NOV-1997; 97US-0066009P.
XX PR 14-NOV-1997; 97US-0066009P.
XX PA (HUMA-) HUMAN GENOME SCI INC.
XX PI Choi GH, Bailey C, Hromockyj A, Kunsch CA;
XX DR WPI, 2003-089120/08.
XX DR N-PSDB; ABX61664.
XX PT New EF040 polypeptides and polynucleotides from Enterococcus faecalis,
XX PT useful for generating an immune response against E. faecalis and other
XX PT Enterococcus species, and as vaccines against other bacterial genera.
XX PS Example 1; Col 115-116; 146pp; English.
XX CC The invention relates to polynucleotide fragments of a gene from
XX CC Enterococcus faecalis, EF040, and the polypeptides encoded by them. The
XX CC polypeptides are useful in detecting E. faecalis, as epitope tags, as
XX CC molecular weight markers on SDS-PAGE gels or for molecular sieve gel
XX CC filtration columns, in generating antibodies that specifically bind to
XX CC the E. faecalis polypeptides, in generating an immune response against E.
XX CC faecalis and other Enterococcus species and as vaccines against other
XX CC bacterial genera. The polynucleotides are useful as probes for gene
XX CC mapping and for identifying E. faecalis in biological samples. Sequences
XX CC ABU13508-ABU13755 represent EF040 polypeptides of the invention. Note:
XX CC The sequence data for this patent can also be obtained from USPTO at
XX CC seqdata.uspto.gov/sequence.html
XX SQ Sequence 560 AA;
SQ Query Match 89.7%; Score 2892; DB 6; Length 560;
Best Local Similarity 100.0%; Pred. No. 1.4e-176;
Matches 560; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 33 BENGSAQVLVHKKKMTDLPPLIONSGKEMSEFDPKYGGLADVTSIYNVTNEFYEQRA 92
DB 1 BENGSAQVLVHKKKMTDLPPLIONSGKEMSEFDPKYGGLADVTSIYNVTNEFYEQRA 60
QY 93 GASVAAKQAVOSLTPGKPVAGTTDANGNVTVOLPKKONGDAVYTIKEBPKEGVAAAT 152
DB 61 GASVAAKQAVOSLTPGKPVAGTTDANGNVTVOLPKKONGDAVYTIKEBPKEGVAAAT 120
QY 153 NNVVAFPVYEMIKQTDGSIKYGTBEIAVVHIIYPRKVVANDSLHYKKGVTAEENGINGAE 212
DB 121 NNVVAFPVYEMIKQTDGSIKYGTBEIAVVHIIYPRKVVANDSLHYKKGVTAEENGINGAE 180
QY 213 PTISSESGSPGVVXYIQYKQDGLYMTTDEKQAKPRITKSKYEIGENDTEAENGNGELT 272
DB 181 PTISSESGSPGVVXYIQYKQDGLYMTTDEKQAKPRITKSKYEIGENDTEAENGNGELT 240
QY 273 VKNLEVGSIYILEEVKAPNNAELIENQKTPPTIEANNQTPVEKTVANDSKYDKTTPSLD 332
DB 241 VKNLEVGSIYILEEVKAPNNAELIENQKTPPTIEANNQTPVEKTVANDSKYDKTTPSLD 300
QY 333 GKDVAIGEKIKYQISVNIPLGIADKEGANKYKFNLVKHDALTFDNTSGEYAVYALY 392
DB 301 GKDVAIGEKIKYQISVNIPLGIADKEGANKYKFNLVKHDALTFDNTSGEYAVYALY 360
QY 393 DDDTVIAPBNVQVTOANGFTVAVNPAYIPTLTPGGTLKFVYFMHINERADPTKQKNA 452
DB 361 DDDTVIAPBNVQVTOANGFTVAVNPAYIPTLTPGGTLKFVYFMHINERADPTKQKNA 420

```

```

QY 453 NVDNHGTDDQTPPEVTEVVGKRFYKVDGDTATATQALAGASFVVRDQNSDPTANYLKIDET 512
DB 421 NVDNHGTDDQTPPEVTEVVGKRFYKVDGDTATATQALAGASFVVRDQNSDPTANYLKIDET 480
QY 513 TKAATWVKTKABATFTTTADGLVDITGLKRGTYYLEEFTVAPDDYVLLTNRIEFVNEQS 572
DB 481 TKAATWVKTKABATFTTTADGLVDITGLKRGTYYLEEFTVAPDDYVLLTNRIEFVNEQS 540
QY 573 YGTTENLVSPKVPKHKGT 592
DB 541 YGTTENLVSPKVPKHKGT 560
RESULT 13
ADY38954
ID ADY38954 standard; protein; 560 AA.
XX AC ADY38954;
XX DT 05-MAY-2005 (first entry)
XX DE Novel Enterococcus faecalis protein sequence SeqID220.
XX KW protein purification; DNA purification; antibacterial; vaccine;
XX KW enterococcus faecalis infection.
XX OS Enterococcus faecalis.
XX PN US2005043528-A1.
XX PD 24-FEB-2005.
XX PF 06-AUG-2004; 2004US-00912362.
XX PR 06-MAY-1997; 97US-0044031P.
XX PR 16-MAY-1997; 97US-004655P.
XX PR 14-NOV-1997; 97US-0066009P.
XX PR 04-MAY-1998; 98US-00071035.
XX PR 29-JUL-2002; 2002US-00206576.
XX PA (HUMA-) HUMAN GENOME SCI INC.
XX PI Choi GH, Bailey C, Hromockyj A, Kunsch CA;
XX DR WPI: 2005-180840/19.
XX DR N-PSDB; ADY38953.
XX PT New isolated Enterococcus faecalis nucleic acid useful for diagnosing,
XX PT preventing and/or attenuating infection by Enterococcus faecalis in
XX PT animals, particularly in humans.
XX PS Claim 9; SEQ ID NO 220; 28pp; English.
XX CC This invention relates to novel genes and the proteins they encode
XX CC isolated from Enterococcus faecalis. The invention may be useful for the
XX CC development of compounds with an antibacterial activity or a vaccine. The
XX CC present invention is useful for the diagnosis, prevention and/or
XX CC attenuation of infection by Enterococcus faecalis in animals,
XX CC particularly in humans. The present sequence is that of a novel
XX CC Enterococcus faecalis protein of the invention. Note: The present
XX CC sequence does not form part of the printed specification but was obtained
XX CC in electronic form from the USPTO web.
XX SQ Sequence 560 AA;
SQ Query Match 89.7%; Score 2892; DB 9; Length 560;
Best Local Similarity 100.0%; Pred. No. 1.4e-176;
Matches 560; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 33 BENGSAQVLVHKKKMTDLPPLIONSGKEMSEFDPKYGGLADVTSIYNVTNEFYEQRA 92
DB 1 BENGSAQVLVHKKKMTDLPPLIONSGKEMSEFDPKYGGLADVTSIYNVTNEFYEQRA 60

```

QY 93 GASVDAKQAVOSLTPEKPVAGSTTDANGNTVYQLEPKONGKDAYTTEKEPEKGVAAAT 152  
DB 61 GASVDAKQAVOSLTPEKPVAGSTTDANGNTVYQLEPKONGKDAYTTEKEPEKGVAAAT 120  
QY 153 NNVAAPVYEMIKQTDGSKYKGTSEELAVVHVPKVVANDGSLHVKVGTAEENGANGAE 212  
DB 121 NNVAAPVYEMIKQTDGSKYKGTSEELAVVHVPKVVANDGSLHVKVGTAEENGANGAE 180  
QY 213 FVYSSESGPVTYKIQGVKDGLYTTTDEKQAKPTGSKYSYIEGNDPTEAENGTELT 272  
DB 181 FVYSSESGPVTYKIQGVKDGLYTTTDEKQAKPTGSKYSYIEGNDPTEAENGTELT 240  
QY 273 VNNLEVGSTYIEBVKAPNNALLENQTKPTTEANNQTPVEKTVKNDTSKYDKTPSLD 332  
DB 241 VNNLEVGSTYIEBVKAPNNALLENQTKPTTEANNQTPVEKTVKNDTSKYDKTPSLD 300  
QY 333 GQDVAIGEKIKYQISVNIPLGIADKEGDKANKYKFNLYDKHDAALTFDNTSGEYVAYLY 392  
DB 301 GQDVAIGEKIKYQISVNIPLGIADKEGDKANKYKFNLYDKHDAALTFDNTSGEYVAYLY 360  
QY 393 GQDVAIGEKIKYQISVNIPLGIADKEGDKANKYKFNLYDKHDAALTFDNTSGEYVAYLY 452  
DB 361 GQDVAIGEKIKYQISVNIPLGIADKEGDKANKYKFNLYDKHDAALTFDNTSGEYVAYLY 420  
QY 453 NVNDGHTDQTPPTVEVTVGKRFKRVDDVATATQALAGSPVVRQNSDPTANYLKIDET 512  
DB 421 NVNDGHTDQTPPTVEVTVGKRFKRVDDVATATQALAGSPVVRQNSDPTANYLKIDET 480  
QY 513 TKAATVTKKATATFTTTADGLVDITGLKYGTYLEETVAPDDVYLTNRLEFVNEOS 572  
DB 481 TKAATVTKKATATFTTTADGLVDITGLKYGTYLEETVAPDDVYLTNRLEFVNEOS 540  
QY 573 YGTENLVSPKVPNNHKGT 592  
DB 541 YGTENLVSPKVPNNHKGT 560  
RESULT 14  
AD084895  
ID AD084895 standard; protein; 625 AA.  
AC AD084895;  
XX  
DT 29-JUL-2004 (first entry)  
XX  
DE E faecium surface anchored LPXTG protein Segid19.  
XX  
KM LPXTG; cell wall-anchored surface protein; Gram positive bacterium;  
KM extracellular matrix molecule; sequence database; C-terminal;  
KM immunoglobulin-like fold region; Ig-like fold region; antibacterial;  
KM vaccine; gene therapy; infection; medical device; prosthesis;  
XX premature newborn; AIDS; debilitated cancer; bone marrow transplantation.  
OS Enterococcus faecium.  
XX  
PN WO2004025416-A2.  
XX  
PD 25-MAR-2004.  
XX  
PF 15-SEP-2003; 2003WO-US028789.  
XX  
PR 13-SEP-2002; 2002US-0410303P.  
XX  
PA (TEXA) UNIV TEXAS A & M SYSTEM.  
PA (INHI-) INHIBITEX INC.  
PA (UABR-) UAB RES FOUNDD.  
XX  
PI Hook M, Xu Y, Sillanpaa JV, Schanman N, Ponnuraj K, Patti JM,  
PI Hutchins JT, Hall A;  
XX  
DR WPI; 2004-315684/29.  
DR N-PSDB; AD084894.

XX  
PT Identifying LPXTG-containing cell wall-anchored surface proteins from  
PT Gram positive bacteria, for treating infection caused by the bacteria,  
PT comprises searching sequence information database for the sequence having  
PT LPXTG-motif.  
XX  
PS Claim 16; SEQ ID NO 19; 96pp; English.  
XX  
CC This invention relates to a novel method of identifying LPXTG-containing  
CC cell wall-anchored surface proteins from Gram positive bacteria that bind  
CC to an extracellular matrix molecule which comprises searching a database  
CC of sequence information for a putative protein sequence having the LPXTG-  
CC motif in its C-terminal region and analysing the sequence for the  
CC presence of one or more immunoglobulin (Ig)-like fold regions. The  
CC invention may be useful for the production of compounds with an  
CC antibacterial activity or for production of a vaccine. In addition the  
CC disclosed sequences may be useful for gene therapy. The antibody is  
CC useful for treating or preventing an infection of Gram-positive bacteria  
CC in a human or animal patient. The method and the proteins are useful in  
CC generating antibodies for treating and preventing the spread of  
CC infections of Gram positive bacteria, for interfering with, or inhibiting  
CC binding interactions by Gram positive bacteria, for monitoring the level  
CC of gram positive bacterial antigens, or antibodies recognising the  
CC antigens in a human or animal patient suspected of containing the  
CC antigens or antibodies, in preventing or reducing infection of medical  
CC devices and prostheses caused by such organisms, and in treating or  
CC preventing infections in highly susceptible groups such as premature  
CC newborns, AIDS and debilitated cancer patients, and bone marrow  
CC transplantation. The present sequence is that of a surface anchored LPXTG  
CC protein identified using the method of the invention.  
XX  
SQ Sequence 625 AA;  
Query Match 75.8%; Score 2446; DB 8; Length 625;  
Best Local Similarity 74.2%; Pred. No. 5,9e-148;  
Matches 462; Conservative 64; Mismatches 95; Indels 2; Gaps 1;  
QY 1 MKQKKWYTVSTLLILPLFTSVLGTTTAPAEENGESAKVYHKKKMTDLPDPLIONSG 60  
DB 1 MKQKKWYTVSTLLILPLFTSVLGTTTAPAEENGESAKVYHKKKMTDLPDPLIONSG 58  
QY 61 KEMSEFDKYQGLADYFYSIVYNTNEFEYEQRAAGASVDAKQAVOSLTPEKPVAGSTTDAN 120  
DB 59 KEMSEFDKYQGLADYFYSIVYNTNEFEYEQRAAGASVDAKQAVOSLTPEKPVAGSTTDAN 118  
QY 121 GNVTVQLEPKQNGKAVYTIKEPEKGVAAATNNVAPVYEMIKQTDGSKYKGTSEELAV 180  
DB 119 GNVTVQLEPKQNGKAVYTIKEPEKGVAAATNNVAPVYEMIKQTDGSKYKGTSEELAV 178  
QY 181 VHIYKQVNVANDGSLHVKVGVGAENGANGAEFVYSKSGSGGTATYQGVNDGLYTTWT 240  
DB 179 VHIYKQVNVANDGSLHVKVGVGAENGANGAEFVYSKSGSGGTATYQGVNDGLYTTWT 238  
QY 241 DKEQAKRPTGSKYSYIEGNDPTEAENGTELTVKNLEVGSTYIEBVKAPNNALLENQTK 300  
DB 239 DKEQAKRPTGSKYSYIEGNDPTEAENGTELTVKNLEVGSTYIEBVKAPNNALLENQTK 298  
QY 301 TPFTTEANNQTPVEKTVKNDTSKYDKTPSLDQKQVAGIKYQISVNIPLGIADKEGD 360  
DB 299 TPFTTEANNQTPVEKTVKNDTSKYDKTPSLDQKQVAGIKYQISVNIPLGIADKEGD 358  
QY 361 ANKYKFNLYDKHDAALTFDNTSGEYVAYLYDQGVTAPENVQVTEQANGSTVAVNPAY 420  
DB 359 ANKYKFNLYDKHDAALTFDNTSGEYVAYLYDQGVTAPENVQVTEQANGSTVAVNPAY 418  
QY 421 IPTLTGGTLKFEVYFMHLEKADPTKGFKNKNAVNDGHTDQTPPTVEVTVGKRFKRVDD 480  
DB 419 IPTLTGGTLKFEVYFMHLEKADPTKGFKNKNAVNDGHTDQTPPTVEVTVGKRFKRVDD 478  
QY 481 GQVTAIGALAGSPVVRQNSDPTANYLKIDETTKAATVTKKATATFTTTADGLVDITG 540  
DB 479 GQVTAIGALAGSPVVRQNSDPTANYLKIDETTKAATVTKKATATFTTTADGLVDITG 538





GenCore version 5.1.9  
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: July 6, 2006, 02:31:46 ; Search time 42 Seconds  
(without alignments)  
1436.379 Million cell updates/sec

Title: US-10-661-809A-13

Perfect score: 3225

Sequence: 1 MKQLKKVYVYVSTLLILPL.....GAVLLIAGVYFARRRKENA 627

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: p1r1:\*  
2: p1r2:\*  
3: p1r3:\*  
4: p1r4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score | Query Match | Length | ID      | Description         |
|------------|-------|-------------|--------|---------|---------------------|
| 1          | 458   | 14.2        | 507    | SS52348 | hypothetical prote  |
| 2          | 296.5 | 9.2         | 665    | P95053  | cell wall surface   |
| 3          | 296   | 9.2         | 1530   | AB1396  | peptidoglycan anch  |
| 4          | 267   | 8.3         | 1612   | AB1347  | probable peptidogl  |
| 5          | 250.5 | 7.8         | 1622   | AB1717  | probable cell surf  |
| 6          | 247   | 7.7         | 1806   | AP1717  | probable peptidogl  |
| 7          | 244.5 | 7.6         | 534    | AB1347  | fimbrial protein t  |
| 8          | 242.5 | 7.5         | 793    | AB1094  | probable peptidogl  |
| 9          | 234.5 | 7.3         | 1816   | P83901  | hypothetical prote  |
| 10         | 229   | 7.1         | 681    | AC1458  | surface anchored p  |
| 11         | 228   | 7.1         | 1983   | G86643  | hypothetical prote  |
| 12         | 220.5 | 6.8         | 537    | A35400  | surface protein T6  |
| 13         | 216.5 | 6.7         | 1661   | AB3695  | hypothetical prote  |
| 14         | 215   | 6.7         | 614    | P86719  | hypothetical prote  |
| 15         | 212.5 | 6.6         | 940    | AB1744  | internalin protein  |
| 16         | 202.5 | 6.3         | 393    | G95053  | cell wall surface   |
| 17         | 196.5 | 6.1         | 533    | A35259  | type 1 fimbrial pr  |
| 18         | 195.5 | 6.1         | 1873   | T30944  | surface protein pr  |
| 19         | 193.5 | 6.0         | 940    | AD1374  | internalin protein  |
| 20         | 192.5 | 6.0         | 1039   | T30856  | protein P2 - Strept |
| 21         | 190.5 | 5.9         | 586    | AD1458  | probable peptidogl  |
| 22         | 189   | 5.9         | 5291   | P90696  | hypothetical prote  |
| 23         | 186.5 | 5.8         | 571    | AI1094  | probable peptidogl  |
| 24         | 185   | 5.7         | 2551   | B98047  | hypothetical prote  |
| 25         | 182   | 5.6         | 5188   | B85547  | probable rtx fam1   |
| 26         | 180   | 5.6         | 745    | S06992  | wall-associated pr  |
| 27         | 180   | 5.6         | 745    | B33856  | hypothetical 80k p  |
| 28         | 178   | 5.5         | 893    | B95053  | cell wall surface   |
| 29         | 177.5 | 5.5         | 1092   | S42798  | fibronectin-bindin  |

|    |       |     |       |   |        |                     |
|----|-------|-----|-------|---|--------|---------------------|
| 30 | 170   | 5.3 | 1939  | 2 | D97316 | probable S-layer p  |
| 31 | 169.5 | 5.3 | 1386  | 2 | AC1533 | surface protein (L  |
| 32 | 169.5 | 5.3 | 1417  | 2 | H90670 | probable invasin l  |
| 33 | 169.5 | 5.3 | 1417  | 2 | D85521 | probable adhesin e  |
| 34 | 169.5 | 5.3 | 4936  | 2 | AB2515 | hypothetical prote  |
| 35 | 168   | 5.2 | 13055 | 2 | T16580 | hypothetical prote  |
| 36 | 167   | 5.2 | 562   | 2 | AB1335 | probable peptidogl  |
| 37 | 166   | 5.1 | 2059  | 2 | D82671 | surface protein Xf  |
| 38 | 165.5 | 5.1 | 814   | 2 | I40048 | S-layer protein pr  |
| 39 | 165.5 | 5.1 | 1185  | 2 | A42404 | collagen adhesin -  |
| 40 | 164.5 | 5.1 | 593   | 2 | S00128 | protein G precursor |
| 41 | 164.5 | 5.1 | 1092  | 2 | T30214 | fibritogen-binding  |
| 42 | 164.5 | 5.1 | 1483  | 2 | C97012 | probably celluloso  |
| 43 | 164.5 | 5.1 | 1684  | 2 | JW0057 | gravin - human      |
| 44 | 164   | 5.1 | 1965  | 2 | S75200 | fat protein - Syne  |
| 45 | 161   | 5.0 | 664   | 2 | S42574 | streptococcal surf  |

## ALIGNMENTS

## RESULT 1

SS52348  
hypothetical protein 2 - Lactobacillus leichmannii  
C:Species: Lactobacillus leichmannii  
C:Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 09-Jul-2004  
C:Accession: SS52348

R:Schenk-Groeninger, R.  
submitted to the EMBL Data Library, January 1995

A:Reference number: SS52347

A:Accession: SS52348

A:Molecule type: DNA

A:Residues: 1-507 <SCH>

A:Cross-references: UNIPROT:Q48707; UNIPARC:UP100000B2PDE; EMBL:X01869; NID:G666067; PIL

A:Experimental source: DSM 20076

Query Match 14.2%; Score 458; DB 2; Length 507;

Best Local Similarity 24.9%; Pred. No. 5.6e-18;

Matches 166; Conservative 81; Mismatches 197; Indels 224; Gaps 26;

|    |     |                                                              |     |
|----|-----|--------------------------------------------------------------|-----|
| QY | 11  | VSTLLILPLFTSVLGT--TTAPAEENGSAQIVHKKKTDLDPDLQNSGKMSFDTK       | 69  |
| DB | 9   | ITGLIMILPLILSLPATPKVSAADTNNSVKTLH-KRVFDSAQBAKQVGTGLMDFG-     | 66  |
| QY | 70  | QGLADVTSPINVTBFEQRAAGASVDAKAYOS-----LTPGKPAVAGTT---          | 117 |
| DB | 67  | TGLNGVTFKAVNVTVHLSLRKSGSAODAVTAIGSDAKSDNLPYAGSAIATETTATS     | 126 |
| QY | 118 | -----DANGNVTVQLPKQNGKDAVTTIKBPKEGVAAATNVVAPVYE               | 162 |
| DB | 127 | KEDDGIAPNDNLKDSQNVQTYL-----FVETDSDPTDTQQAAPVLTWPIY-          | 175 |
| QY | 163 | MIKQTDGSKYKGTBELAVH---ITPKNVVANDGSLHVKKVGTAENEGANGAEFVYSKE   | 219 |
| DB | 176 | --KTSPTS-----AIHNDIQIYKVN-----                               | 194 |
| QY | 220 | GSPPGVYKIQGVKDGIVTTTIDKEQAKRITTSKSYEIGENDPTAEANGTGLTVKNLBYG  | 279 |
| DB | 195 | -----                                                        | 194 |
| QY | 280 | SYLSEVKAPNNAEILNQTKPTTEIANNQTPVEKTVNDTSKYDKTTPSIDGKV---      | 336 |
| DB | 195 | -----KSPITFDL--DEASKDLAVTLPLDGSRTVNA                         | 224 |
| QY | 337 | AIGKIKIQISVINPLGADKEGDANKYKFLVLDKHAALFPDNTVSEIYALVDSOT       | 396 |
| DB | 225 | QYKSGFGVYITNVVWNKDKD-----TFNVVVKPDGIDIDIST-----VSIDGLT       | 271 |
| QY | 397 | VIAPENVOTBEQANGTVA--VNPAVITPLTFGGTLKFYVFMHLEKADPTGKFKREA--   | 452 |
| DB | 272 | --KSTDYVNNKONGVQVFFKTTSAVQALA-GKSLITTYKATLTNNALPDDAIGNTATL   | 328 |
| QY | 453 | NVDNGHTDDQTPPT-VEVVTGGRFLKYDGDVATATQALAGASF--VVRDQNSDTANYLKI | 509 |

Db 329 SIGGNITITSPANGPRIYTCGAQFVKDSQ--SNKTLIAGAEFQLVKVDNSNGNIVSY--- 383  
 Qy 510 DETKAA-----TWKTKAEATFTTTADGLVDIGLKG-----TYLLEETVAPDDYV 558  
 Db 384 --ATQASGSYTWMDSATBEATLYTSDANGLVALNGLSTSDKLDSGESYALLEIOAPGGYA 441  
 Qy 559 LITRIEIVVNAEGSVGTENLVSPKVPNKHKGLPSTGCGIIVYLGSNAVLLIIA-GV 617  
 Db 442 KLDSVVKESIQGSFGDSNKIT---IDNTEGLLPSTGCGIYIFLAIGIVIMIVAFGG 497  
 Qy 618 YFARRRKE 625  
 Db 498 YKAIKKHE 505

## RESULT 2

cell wall surface anchor family protein [imported] - Streptococcus pneumoniae (strain T10000005144B; GB:AE005672; PIDN:AAK74623.1; C:Species: Streptococcus pneumoniae C:Date: 03-Aug-2001 #sequence\_revision 03-Aug-2001 #text\_change 09-Jul-2004 C:Accession: F95053 R:Retceltin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Heidison, J.D.; Unanue, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapple, nson, T.; Hickey, E.K.; Holt, I.E. Science 293, 498-506, 2001 A:Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison, A:Title: Complete Genome Sequence of a Virulent Isolate of Streptococcus pneumoniae. A:Reference number: A95000; MUID:21357209; PMID:11463916 A:Accession: F95053 A:Status: preliminary A:Molecule type: DNA A:Residues: 1,665 <KUP> A:Cross-references: UNIPROT:O97SC2; UNIPARC:UPI000005144B; GB:AE005672; PIDN:AAK74623.1; A:Experimental source: strain T10000005144B; C:Genetics: A:Gene: SP0463

Query Match 9.28; Score 296.5; DB 2; Length 665;

Matches 176; Conservative 84; Mismatches 259; Indels 247; Gaps 38;

QY 1 MKQLKQWYTVSTLLILPLFTSVLGTTTAFABENGESAQLVIHKKMTDLPDPLIQNSG 60

Db 1 MKSINKFLMLALLTA--SSLSFSAATVFA-AGTTTTSVTVHKLATDGMMDKIANE- 55

61 KEMSEFDKYOG-----LADVTFSIYNTNEFYEO--RAGASVDAKAQAVOSL 106

Db 56 ---LETGNYAGNKVGLPANAKEIAGVMFWTNTNNEIIDENGOTLGVIDPOTFKLSGA 112

107 TPBKPVACGTTDANGNV--TVOLPKKONGDAVYTIKEBPKEGVAAATNMVVAEPVYEMI 164

Db 113 MPATAMK-I-TEAECAKENTANT.P-----AAKYKIYEITH 145

04 165 KOT-----DGSYKCTEER AV-----YHTVPRNF-----YANDGSLH 196

[illegible]

כך הנהיגו את המלחמה, וזו היתה הסיבה שכל המלחמה הזאת לא היתה מלחמה של אומות, אלא מלחמה של אנשים.

[illegible]

1

A vertical strip of a film negative, showing horizontal lines and some markings. The markings include two pairs of dots and two pairs of horizontal lines.

[illegible][illegible][illegible]

27 37 47 57 67 77 87 97 107 117 127 137 147 157 167 177 187 197 207 217 227 237 247 257 267 277 287 297 307 317 327 337 347 357 367 377 387 397 407 417 427 437 447 457 467 477 487 497 507 517 527 537 547 557 567 577 587 597 607 617 627 637 647 657 667 677 687 697 707 717 727 737 747 757 767 777 787 797 807 817 827 837 847 857 867 877 887 897 907 917 927 937 947 957 967 977 987 997 1007 1017 1027 1037 1047 1057 1067 1077 1087 1097 1107 1117 1127 1137 1147 1157 1167 1177 1187 1197 1207 1217 1227 1237 1247 1257 1267 1277 1287 1297 1307 1317 1327 1337 1347 1357 1367 1377 1387 1397 1407 1417 1427 1437 1447 1457 1467 1477 1487 1497 1507 1517 1527 1537 1547 1557 1567 1577 1587 1597 1607 1617 1627 1637 1647 1657 1667 1677 1687 1697 1707 1717 1727 1737 1747 1757 1767 1777 1787 1797 1807 1817 1827 1837 1847 1857 1867 1877 1887 1897 1907 1917 1927 1937 1947 1957 1967 1977 1987 1997 2007 2017 2027 2037 2047 2057 2067 2077 2087 2097 2107 2117 2127 2137 2147 2157 2167 2177 2187 2197 2207 2217 2227 2237 2247 2257 2267 2277 2287 2297 2307 2317 2327 2337 2347 2357 2367 2377 2387 2397 2407 2417 2427 2437 2447 2457 2467 2477 2487 2497 2507 2517 2527 2537 2547 2557 2567 2577 2587 2597 2607 2617 2627 2637 2647 2657 2667 2677 2687 2697 2707 2717 2727 2737 2747 2757 2767 2777 2787 2797 2807 2817 2827 2837 2847 2857 2867 2877 2887 2897 2907 2917 2927 2937 2947 2957 2967 2977 2987 2997 3007 3017 3027 3037 3047 3057 3067 3077 3087 3097 3107 3117 3127 3137 3147 3157 3167 3177 3187 3197 3207 3217 3227 3237 3247 3257 3267 3277 3287 3297 3307 3317 3327 3337 3347 3357 3367 3377 3387 3397 3407 3417 3427 3437 3447 3457 3467 3477 3487 3497 3507 3517 3527 3537 3547 3557 3567 3577 3587 3597 3607 3617 3627 3637 3647 3657 3667 3677 3687 3697 3707 3717 3727 3737 3747 3757 3767 3777 3787 3797 3807 3817 3827 3837 3847 3857 3867 3877 3887 3897 3907 3917 3927 3937 3947 3957 3967 3977 3987 3997 4007 4017 4027 4037 4047 4057 4067 4077 4087 4097 4107 4117 4127 4137 4147 4157 4167 4177 4187 4197 4207 4217 4227 4237 4247 4257 4267 4277 4287 4297 4307 4317 4327 4337 4347 4357 4367 4377 4387 4397 4407 4417 4427 4437 4447 4457 4467 4477 4487 4497 4507 4517 4527 4537 4547 4557 4567 4577 4587 4597 4607 4617 4627 4637 4647 4657 4667 4677 4687 4697 4707 4717 4727 4737 4747 4757 4767 4777 4787 4797 4807 4817 4827 4837 4847 4857 4867 4877 4887 4897 4907 4917 4927 4937 4947 4957 4967 4977 4987 4997 5007 5017 5027 5037 5047 5057 5067 5077 5087 5097 5107 5117 5127 5137 5147 5157 5167 5177 5187 5197 5207 5217 5227 5237 5247 5257 5267 5277 5287 5297 5307 5317 5327 5337 5347 5357 5367 5377 5387 5397 5407 5417 5427 5437 5447 5457 5467 5477 5487 5497 5507 5517 5527 5537 5547 5557 5567 5577 5587 5597 5607 5617 5627 5637 5647 5657 5667 5677 5687 5697 5707 5717 5727 5737 5747 5757 5767 5777 5787 5797 5807 5817 5827 5837 5847 5857 5867 5877 5887 5897 5907 5917 5927 5937 5947 5957 5967 5977 5987 5997 6007 6017 6027 6037 6047 6057 6067 6077 6087 6097 6107 6117 6127 6137 6147 6157 6167 6177 6187 6197 6207 6217 6227 6237 6247 6257 6267 6277 6287 6297 6307 6317 6327 6337 6347 6357 6367 6377 6387 6397 6407 6417 6427 6437 6447 6457 6467 6477 6487 6497 6507 6517 6527 6537 6547 6557 6567 6577 6587 6597 6607 6617 6627 6637 6647 6657 6667 6677 6687 6697 6707 6717 6727 6737 6747 6757 6767 6777 6787 6797 6807 6817 6827 6837 6847 6857 6867 6877 6887 6897 6907 6917 6927 6937 6947 6957 6967 6977 6987 6997 7007 7017 7027 7037 7047 7057 7067 7077 7087 7097 7107 7117 7127 7137 7147 7157 7167 7177 7187 7197 7207 7217 7227 7237 7247 7257 7267 7277 7287 7297 7307 7317 7327 7337 7347 7357 7367 7377 7387 7397 7407 7417 7427 7437 7447 7457 7467 7477 7487 7497 7507 7517 7527 7537 7547 7557 7567 7577 7587 7597 7607 7617 7627 7637 7647 7657 7667 7677 7687 7697 7707 7717 7727 7737 7747 7757 7767 7777 7787 7797 7807 7817 7827 7837 7847 7857 7867 7877 7887 7897 7907 7917 7927 7937 7947 7957 7967 7977 7987 7997 8007 8017 8027 8037 8047 8057 8067 8077 8087 8097 8107 8117 8127 8137 8147 8157 8167 8177 8187 8197 8207 8217 8227 8237 8247 8257 8267 8277 8287 8297 8307 8317 8327 8337 8347 8357 8367 8377 8387 8397 8407 8417

DD 358 P1FAG-----AEEAIFDLVNAGIGKVQIV-----ILILIDNN---I 389

```

QY      407 EQANGFYVAANPAIIPLLTPOGTLKFFYFPHLEKNDPTGKFKHAAVNDGHTDDOTPT 466
Db      390 VTVAAGLKRNTPEYKVERSISKGSADYOEITTAGIA--VNKKODE----NPKPLDPTBP- 442
QY      467 VEAVTGGKRFKIV--DGCVTATQALAGASFVVRDQNSDPA--NYL--KIDE-----T 512
Db      443 -KVLTGKKFKVNDKO-----NRLAGABFVL--ANADNAGGYLARKADKVSQEBKQLVIT 495
QY      513 TK-----AATWV--KTAETTF 528
Db      496 TKDALDBRAVAAYNALTAQOOTOQEKEKVDKAGAAAYNAAYTAANNAFEMVADKDNENVKL 555
QY      529 TTTADGLVDITGLKKGYVYLEETVAPDDVYLLTRIEFVNEOSGYTTEILV----- 580
Db      556 VSDAGRFETIGLAGTYLEETKQAPAGYALLTSRQKFBFVTAYSYATNGGIEYTAGSK 615
QY      581 -SPKVPNKHGKTLPTSGKGIVYLLSGAVLLLIAGVTPARRKKE 625
Db      616 DDATKRVNK--KITIPOTGGIGITLIFAAVAGAIMIAYAVAKNKD 660

```

### RESULT 3

AHI336  
 peptidoglycan anchored protein (LPXTG motif) [imported] - *Listeria monocytogenes* (strain  
 C)Species: *Listeria monocytogenes*  
 C.Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #ext\_change 09-Jul-2004  
 C.Accession: AHI396  
 R.Glasser, P.; Frangoul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloekerer,  
 J.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Duseurget, O.; Entian, K.D.; Feihl, H.  
 D.; Jones, L.M.; Karet, U.  
 Science 294, 849-852, 2001  
 A.Authors: Krett, U.; Kunh, M.; Kunst, F.; Kurapkat, G.; Madeno, E.; Maitournam, A.; Ma  
 ok, C.; Schuener, T.; Simoes, N.; Terez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,  
 A.Title: Comparative genomics of *Listeria* species.  
 A.Reference number: AB1077; MUID:21537279; PMID:11679669  
 A.Accession: AHI396  
 A.Status: preliminary  
 A.Molecule type: DNA  
 A.Residues: 1-1530 <GLA>  
 A.Cross-references: UNIPROT:Q8Y479, UNIPARC:UPF010000054CB4, GB:NC\_003210, PTD:CAD00654.1  
 A.Experimental source: strain EGD-e  
 C.Genetics:

A; Serie: 11023/6

Query Match 5.28; Score 250; DB 2; dengcm 1550  
Best Local Similarity 22.2%; Pred. No. 2.6e-08;

|         |     |              |    |            |     |         |      |      |     |
|---------|-----|--------------|----|------------|-----|---------|------|------|-----|
| Matches | 16/ | conservative | 1/ | mismatches | 23/ | inserts | 2/0/ | gaps | 30/ |
|---------|-----|--------------|----|------------|-----|---------|------|------|-----|

28 T T A F A E N G E S A Q L V I H K K M I D L P D P L I Q N S G N E M S E F D K I Q G L A D V I F S I I N V I N E F I 8'

D5 876 TIGDGTNGEITGKILNK---VDKADPSIFLEG-----AIFDLY----- 911

QY 88 EQRAGASVDAKQAVOSLTPGKPAQGTTDANGNVIVQLPKQNGKDAV---YTIKE-E 143

D6 912 ---ANDEKVDTQ-----TTDKNGVIEFD-----DLVYGDYTLKEVS 944

QY 144 PKEGVVAATNMVAFPVYEMIKQTDGSYKYGTEELAVHIYPKNVANDGSLHVKKVGTA 203

Db 945 APBGYTLPTASTENIQ-----KLEQDEKVVQVMNEKMPICKETGEVHLVKTDKA 993

QY 204 ENGLNGAEFVISKSESP-----GTVKYLQGVKDGLYTW 238

Db 994 TGATLGAERFSLYDKSGAELQNGLTTDENGELTIHNLDSYLLKETKAPEGYKLSEKTW 1053

QY 239 TTDKEQ-----AKRFITGKSYEIGENDFTEA 264

Db 1054 EFSVESGQVDAIEIQAENEKDLGEAVLTQVDSETNAKLSGAKENLLNDSGEVIQTNLVSD 1113

QY 265 ENGTELTVKNLEVGSIILEBVKAPNNAELIENQTPTPTIEANNQTPVEKTVKND-TSK 323

Db 1114 EN--GEIRVQNLBPGDYAFQETEAFTNYDLATN--TWPFITIVAGQTSATMTVAENNKTK 1165

```
QY 324 VDKTPS--LDKQVAIGEKIKYQISVNIPLGIADKEGDANKYKFNLYDKHDAALTFDN 381
DB 1170 PVDVTEBVLIVKQDSATGETLEGAV---FDLMTRD-----GAIVASNLTTDANGELTVN 1221
QY 382 VTSGEYVALYDGDYIAPENYQVTEQANGFTVAVNPAYIPILJ-----PGGTLYK 431
DB 1222 LAFGKYSF-----KETKABEGVELADVMEFTIAPNQPEKITTAENTKLAIPIDAGSVK 1276
QY 432 FYVFMHLNKAADPTGFK-----K-----NE 451
DB 1277 II-----KQSENGVRLAGAFSLIANGETLQNLTKTDEAGELEVNNLBGNRYIOE 1329
QY 452 ANVDNGHTDDQPTPEVYV-----TGKRPFIKVDGVTATQAL 489
DB 1330 TKAPDQVQESTWQGEIYVANDTSQVTVIAENAKLEBDVAETGAVLLIKTDSF-TGTR-L 1387
QY 490 AGASPVVRDQNSDPTANYLKIDETTKATWVKTAEMTPTTTADGLVDITGLKRYGYLYE 549
DB 1388 SGAVFSLDDESG-----KVIOANLTTDENGELFIDGLTPGNVSLK 1427
QY 550 ETYAPBDYVL-----LTKRIEFVY--NEOSYGTETLVLS-----PR--KVPNK 588
DB 1428 ETKAPDGYELAQPMNFOIVKQGVDAVILKAENSPILANGALISFEGDETDKPSISIRI 1487
QY 589 HKGT-----LPSTGAKGYV--VYLGSAVLL 612
DB 1488 KTDITLATEVTKLPQTGDKTSFPRVILGGSATVLM 1520
```

## RESULT 4

```
AB1347
probable pepitidoglycan bound protein (LPXTG motif) lmo2178 [imported] - Listeria monocyt
C/Species: Listeria monocytogenes
C/Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #ext_change 09-Jul-2004
C/Accession: AB1347
R/Glaeser, P.; Frangoul, L.; Buchtleier, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Duseurget, O.; Entian, K.D.; Feihl, H
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A/Authors: Krefz, J.; Kuhn, M.; Kunst, F.; Kurapkac, G.; Madueno, E.; Maltournam, A.; M
ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlend,
A./Title: Comparative genomics of Listeria species.
A/Reference number: AB1077; MUID:21537279; PMID:11679669
A/Accession: AB1347
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-1612 <GLA>
A/Cross-references: UNIPROT:Q8Y591; UNIPARC:UPI00000CCF1B2; GB:NC_003210; PIDN:CAD00256.1
A/Experimental source: strain EGD-e
C/Genetics:
A/Gene: lmo2178
```

```
Query Match 8.3%; Score 267; DB 2; Length 1612;
Best Local Similarity 21.5%; Pred. No. 1.1e-06;
Matches 145; Conservative 88; Mismatches 226; Indels 214; Gaps 26;

QY 35 NGEASQVLIHKKKMDLP-----PLIQSGKSESEF-----DKYGLADVTFSITVY 82
DB 888 NNIISADRPINNPITITDTMTKGTGAGVXSSFCVNMETTGEDIDSKTY---DITVYDNF 944
QY 83 TNEFYEQRAAG-----ASVD-----AAKQAVOSLTPGKPVQ 114
DB 945 TIQFDYKATAPIKYTYSTISLSGLVSNATTAASDYSGLPMTYKSRSTTSISAPATIS 1004
QY 115 GTTDAN-GNVTY-QLPKQNGKD-----AVYTIKEPREGVVAATN---MVAAPV--- 160
DB 1005 GSGATIGSLIKITKVKDKNSKLTGAKPOLYTL-BGDKAGGATYDSDGKIYMDGIQSG 1063
QY 161 -YEMIK-----QTDGSKYKGTTELAVVHIYKNNVAN---DGLSHVKKVNGAHEGNA 209
DB 1064 KTKLVETAPGTYSIDSEKDGKEITVTADTATSVTIENTQGSVLLKEBSATYDADA 1123
QY 210 GAEFVSKSEG-----SPGTAVKIYQ----- 229
```

```
DB 1124 GAEFELQNDGTVADNLVSNADGKLEVTDLARGDYQFVETKAPTVGLDGAFTKPTVER 1183
QY 230 -----GVKGLYTTWTDKQAKRFITGKSYEIGENDFTAE-----NGGELT 272
DB 1184 NOEAAVIVTKENTAKGSVVLTKEDSFKATISGAEEFLQNAVTGKCNLTNNAGKLE 1243
QY 273 VKNLVGSYILBEEKAPNNABLIENOTKPTFEANNQTPVEKTVKNDTSKVDKTPSLD 332
DB 1244 VTDLAPGDYKFEVETKAPTVGEL--DAPVTFIIEFNQATVANNVTKENATGSVVLTKLD 1301
QY 333 GKQVAIGEKIKYQISVNIPLGIADKEGDANKYKFNLYDKHDAALTFDNTSGEYVALY 392
DB 1302 SKRSNLAGEFELQKGLVSLKDK-----VTEANGQLIDN----- 1339
QY 393 DGDYIAPENYQV--TEQANGFTVAVNPAYIPILTTGGLTKFYVFMHLNKAADPTGPKN 450
DB 1340 ----LAPGDYQVLEKAPTVGDLDTATPVEF-----TIEF-----NOKA----- 1373
QY 451 EAVNDNGHTDDQPTPEVYVVTGSKRFIKVDGVTATQALAGASPVVRDQNSDPTANYLKID 510
DB 1374 -----PIQVTKTNSTGVSVLTKTDETKA--LLANATFKLVDEBNVTENL--- 1419
QY 511 ETKATWVKTKAEATFTTTADGLVDITGLKRYGYLYEBETVAPDDYVLLT---NRIF 566
DB 1420 -----TTDASGKLEITNLAPGQVLETKAPAGVBLDPTVAVDVKITF 1461
QY 567 VNEOSYGTETLV 579
DB 1462 DQKETLQVTKTNL 1474
```

## RESULT 5

```
AB1717
probable cell surface protein (LPXTG motif) [imported] - Listeria innocua (strain Clijf
C/Species: Listeria innocua
C/Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #ext_change 09-Jul-2004
C/Accession: AB1717
R/Glaeser, P.; Frangoul, L.; Buchtleier, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Duseurget, O.; Entian, K.D.; Feihl, H
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A/Authors: Krefz, J.; Kuhn, M.; Kunst, F.; Kurapkac, G.; Madueno, E.; Maltournam, A.; M
ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlend,
A./Title: Comparative genomics of Listeria species.
A/Reference number: AB1077; MUID:21537279; PMID:11679669
A/Accession: AB1717
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-1622 <GLA>
A/Cross-references: UNIPROT:Q92974; UNIPARC:UPI00000CC7F5; GB:AL592022; PIDN:CAC97509.1
A/Experimental source: strain Clijf1262
C/Genetics:
A/Gene: lin2281
```

```
Query Match 7.8%; Score 250.5; DB 2; Length 1622;
Best Local Similarity 20.7%; Pred. No. 9.4e-06;
Matches 171; Conservative 99; Mismatches 252; Indels 305; Gaps 32;

QY 14 LLLILPLFTSVLGTTAFAEENGSAQVIHKKM--TDLPPPL-----IQNSGKMS- 64
DB 884 LNIITPRLTSLPVTINRLQGNQTAQPIKNSQVINTNNEINENENTDITREGNTFTI 943
QY 65 EFDKTOGLADV-----TFSITN--VTNE-----FYEQRAAGAS----- 95
DB 944 OFKNYTAAPIRKVNSTISLSGPISNETTVEAEDPSNVPMPFKKNVAVSPVFTVGTGS 1003
QY 96 -----VDAKQAVOSLTPGKPVAGOGTTDANGVTVOQLPKQNGK 134
DB 1004 GIATITGTYITVDEDDTTKJLGAKFQLYTLDGKSGGEITNNSGELL--LDGIQSGK 1061
QY 135 DAV-----YTIKEPREGVVAATNNVAVPVEYEMIKQTDGSKYKGTTELAVVHIYK 186
```

Db 1062 YKLVEKTAPEGYNISDEYKEG-----KEITVSSSG-----EELILT---IK 1099  
QY 167 NVNVADGSLHVKVQTAENEGINGAEFVYSKSESGPTVKYIQGVKGLYTWTTDKEQAK 246  
Db 1100 NAM-KKGVNLTAKKQASASHEVLADAEFELQNAFGSK-----LKEKLI---TTD----- 1142  
QY 247 RPTGSKSEYIENGDFTLEANGTGEVLTKNLEVGSYLLEEVKAPNNALLENQTKPTPTIE 306  
Db 1143 -----ASGNIEITDLAPGDYLIETKAPTQGL---DATPVNFTID 1180  
QY 307 ANNOPTVEKTVKNDTSKDKTTPSLDGOVAIGEKIKYQISVNIPLGIADKE--GDANKY 364  
Db 1181 F-NQSEAAKVSKTNTAKTG--TVMLTAKQSA-----TNALADATPELRNEDETL 1227  
QY 365 VKENLVDKHDAALTGDVNTSGEYAAVALYGDVTIAPENYQVTEQANGFVAVNPAVYIPTL 424  
Db 1228 VARENVTDDNGEISVADLAPGDYKLI-----ETKAPTGYQDATTVPNFTIDNQSVEVY 1282  
QY 425 TPGGTLKPYFMHLN-----EKADPTKGFKNENAV---DNGHTDDQPTPEVEVVTGK 474  
Db 1283 SKTNNAKGTAVLTKKDSATNBLADATPELRNESPVLVRENVLVDDNGEISVADLAPG- 1341  
QY 475 RFIKYDGDVTAQTALAG-----ASFVVRDQNSDTANTYKIDETTKAATWYKTAEAAT 527  
Db 1342 -----DYKLIETKAPAGYQDADATPVNFTIDFNQSEAAKVSKin-TAKTGTVLTKKDSAT 1395  
QY 528 FTTTAD-----GLVDITGKYGTYLLEETVAPDDVVLTLNRIE 565  
Db 1396 NAELADATPELRNESPVLVRENVLVDDNGEISVADLAPGDYKLIETKAPTGYQDATTVPN 1455  
QY 566 FVV-----NEQSYGT----- 575  
Db 1456 FTIDFNQSEAAKVTNTKKKIGTIIYNFIDVDGNQNDKSVHTGNVGEYSVKAKEIVGH 1515  
QY 576 -----TEN 578  
Db 1516 TLVKDPANKGMYKETSQETITVEYEKKAMPITVEPTPEPSKTEOLTESATVAEPKIKON 1575  
QY 579 LVSPKVPNKHGKTLPTSGKGIVYVYLGSAVLLLAGVYFAPRRKE 625  
Db 1576 IKTTKSPN-NRKCLPSTGDEPPYTMFTG-LFVSVAAGVFLRKPKQ 1620

RESULT 6  
AF1717  
probable peptidoglycan bound protein (LPXG motif) lin2282 [imported] - *Listeria innocua*  
C:Species: *Listeria innocua*  
C:Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 09-Jul-2004  
C:Accession: AF1717  
R:Glaser, P.; Brangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker  
; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Duesunger, O.; Entlian, K.D.; Fehli, H.  
D.; Jones, L.M.; Karsel, U.  
Science 294, 849-852, 2001  
A:Authors: Kreft, J.; Kuhn, M.; Kunat, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Me  
ok, C.; Schueller, T.; Simoes, N.; Tietzer, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlend,  
A. Title: Comparative Genomics of *Listeria* Species.  
A:Reference number: AB1077; MUID:21537279; PMID:11676696  
A:Accession: AF1717  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-1806 <GLA>  
A:Cross-references: UNIPROT:Q929U3; UNIPARC:UP100000CC7P6; GB:AL592022; PIDN:CAC97510.1;  
A:Experimental source: strain Clp11262  
C:Genetics:  
A:Gene: lin2282

Query Match 7.7%; Score 247; DB 2; Length 1806;  
Best Local Similarity 22.0%; Pred. No. 1.7e-05;  
Matches 135; Conservative 63; Mismatches 219; Indels 198; Gaps 20;  
QY 27 TTTAPAEKGEAGQVLIHKKMTDLDPDLIUNGSKGMS-EFDKYGSLADVTTSIVYVNE 85  
Db 1177 TETKTFVEBNQTAIVTK-----ENTAKAGSVVLTQDSATKATIA---GAE 1221

QY 86 FYEORAAQASY-----DA-AKQAVOSLTPGKPAVQGTTDANGNTVOLPKK-----ONG 133  
Db 1222 PELQKADGKTGSENLVTTBADKGIENVNDLAPGYQGVETKKAAGVLDSPSPFTIEFNQD 1281  
QY 134 KDVAVYTIKEEPKEGV-----AATNVVAFPVY-----MIKQTDG----- 169  
Db 1282 KAAVTKENTAKTGSVVLTKQDSVTKATLAGEFELOKADGKTVSENLVTTDVGKIEVND 1341  
QY 170 -----SYKKGTEBLAVVHIYPKVNVANDGSLHVKYKGTLEN 205  
Db 1342 LAPGYQVETKKAAGVLDTPYKFKVGFNQTSVAAVAKEN-TAKAGSVLTKQDSVTK 1400  
QY 206 EGLNGAEFVYSKSESGPTVKYIQGVKGLYTWTTDKEQAKRPTGKSEYIENDPTEAE 265  
Db 1401 AALSDAEFLQTAATK-----VKONLTT----- 1424  
QY 266 NGTGEITVKNLEVGSYLLEEVKAPNNALLENQTKPTPTIEANNOPTPEKTVKNDTSKVD 325  
Db 1425 NASGEIEVADLAPGDYKKEVETKAPTGYEL--DATPVTETIEFNQATAVKVTENYAKTGS 1482  
QY 326 KTTPELDGKDAVIGKIKYQISVNIPLGIADKEGDANKVYKFNLYDKHDAALTDPDNTSG 385  
Db 1483 VVLTKLDSKSKSNLAGAEFLQTKLGVSUKDK-----VTEANGQLQIDNVLAPG 1531  
QY 386 EYAAVALYGDVTIAPENYQVTEQANGFVAVNPAVYIPTLTPGTLKPYFMHLNEKADPT 445  
Db 1532 DYQLV-----ETKAPTGYELADTPVEFTIEFN-----QKQAV 1563  
QY 446 KQFKNENAVDNGHTDDQPTPEYEVYTGKRFKVDGDTATQALAGASPVVRDQNSDPAN 505  
Db 1564 -----QVTKTKMKSTGSVVLTKTGQTKA--PLADATFKVLADANNVTE 1605  
QY 506 YLKIDETTKAATWYKTAATFTTTADGLVDITGLKGYTYLLEETVAPDDVVLTLNRIE 565  
Db 1606 NL-----TTDASGKLEITNLAPDYOLITKAPKGYELDTVPV 1644  
QY 566 FVNVESQYGTENTLV 580  
Db 1645 VKI-----GFNONQV 1654

RESULT 7  
A32347  
fimbrial protein type 2 precursor - *Actinomyces naeslundii*  
C:Species: *Actinomyces naeslundii*  
C:Date: 21-May-1990 #sequence\_revision 21-May-1990 #text\_change 05-Oct-2004  
C:Accession: A32347  
R:Yeung, M.K.; Cisar, J.O.  
J. Bacteriol. 170, 3803-3809, 1988  
A>Title: Cloning and nucleotide sequence of a gene for *Actinomyces naeslundii* wVU45 type  
A:Reference number: A32347; MUID:88314866; PMID:2900829  
A:Accession: A32347  
A:Molecule type: DNA  
A:Residues: 1-534 <YEU>  
A:Cross-references: UNIPROT:P12616; UNIPARC:UP1000012MA55; GB:M21976; NID:q141847; PIDN:  
C:Superfamily: fimbrial subunit precursor  
C:Keywords: transmembrane protein  
F:1-32/Domain: signal sequence #status predicted <SIG>  
F:33-534/Product: fimbrial protein type 2 #status predicted <MAT>

Query Match 7.6%; Score 244.5; DB 2; Length 534;  
Best Local Similarity 23.4%; Pred. No. 3.9e-06;  
Matches 147; Conservative 63; Mismatches 220; Indels 197; Gaps 29;  
QY 90 RAAGASVDAKQAVOSLTPGKPAVQ-----TTDANGNTVOLPKKQNKDAVY--- 138  
Db 10 RRAAAAGVLTAAVIGLA---PMAQAEANMHDINTALGSLTIH--KRLNDGNPITGAP 64  
QY 139 --TIKEEPKEGVAAATNVVAFPVYEM-----IKQDGSYKGTETELAVVHIYPKVNVAND 192  
Db 65 DGTASNDGKGAPVSG--VQFTAYEINGIDIKTSBGNAK--VNLITNGALPDAACANP 119

```

Qy 193 GSLHYKVVGTAEINLGAEFVYSKSESGPGTVKVIQGVNDGLYTWTTDKEQAKREFTGK 252
Db 120 G-----QPTLPNTFRSSRVSQD-----TDRD-----141
Qy 253 SYEIGENDFTEANGTGLTVKNLEVGSIYILEBKAPNNAELENTKPTPTTEANNQTP 312
Db 142 -----GEAKIESLPVKAYLVCECTKTPGNIV-----QKAPFVVTIIPHPNT 181
Qy 313 VKET-----VKNTSKVDKTTPTSLDGDVAIGEKIKQISVNIPLGIADKEGA 361
Db 182 AAKAGGTMLVDVHVYPKNEKIEVAKTIEDQRNNGYIVGSKVRPVSSTLP-----KLD 235
Qy 362 NKVYF-----NVVDKDALTFDQNTSG-----EYAVLYDQGTVIAPENYQVTEOA 409
Db 236 NSEYKQYQKQDLDNRLLKQVNTADYTLGSTRLEDEGDTLGT-DGQTV-----YTFNQ 288
Qy 410 NGFTVAVNDAYIPTLTGSGTLKFVYFMHLEKAD-----PTKGFKE 451
Db 289 NGLSKLKG-----NPGQKQAVFEGVSEVGDGSIINNTAQILSDTYAQPAPBETPP 341
Qy 452 ANVDNGHTDDQTPPTVEVNT-----GGRFVKVDGDVTA--TQALAGSPVY--RDQNSPT 503
Db 342 ANPDN-----PPTTEQVTSKMGDLTIKKVDGNDRSGDKGLKGAFOIYKAKQAVADT 394
Qy 504 ANYLKIDETTKATWVKTKABATFTTTADGLVDITGL-----KYGTY 546
Db 395 C-----SPENDGQPLTNGESTFTTGCGTINFPALFVSDSVQDGTGRNDVADPHRY 447
Qy 547 YLEETVADDDVYLLTNRIEFVNEQSYGTENLV--SPEKVPNKHGTLPTSGKGIVY 603
Db 448 VLVEKAPAGYVLPADASRAITVEPGAQVTOQVVIDNVKQSVG-----LPLTANGMLI 502
Qy 604 YLGSAGVLLIA--GYVFA--RRREN 626
Db 503 LTASGALLMLAVGSVLVARYRERKEN 529

```

## RESULT 8

AH1094

Probable peptidoglycan bound protein (LPXTG motif) Imc0159 [imported] - *Listeria monocytogenes*  
 C.Species: *Listeria monocytogenes*  
 C.Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 09-Jul-2004  
 C.Accession: AH1094  
 R.Glasser, P.; Frangoul, L.; Buchliesser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker, D.; Jones, L.M.; Karsc, U.  
 Science 294, 849-852, 2001  
 A.Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkac, G.; Madueno, E.; Maltournam, A.; Maok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voess, H.; Wehland, A.  
 A.Title: Comparative genomics of *Listeria species*  
 A.Reference number: AB1077; MUID:21537279; PMID:11679669  
 A.Accession: AH1094  
 A.Status: preliminary  
 A.Molecule type: DNA  
 A.Residues: 1-793 <GLA>  
 A.Cross-references: UNIPROT:O8YAG7; UNIPARC:UPI0000054CA9; GB:NC\_003210; PINN:CA098374.1  
 A.Experimental source: strain EGD-e  
 C.Gene: Imc0159

```

Query Match 7.5%; Score 242.5; DB 2; Length 793;
Best Local Similarity 20.9%; Pred. No. 9, 1e-06;
Matches 177; Conservative 96; Mismatches 286; Indels 287; Gaps 32;

Qy 4 LKKVYVSTLLILPLF--TSVLGTTTFAEENGESAOVTHKKKMTLDPPLIQNSGK 61
Db 8 LKKVGLAFSLIVASTITPQTVVVKATSYSGQLATVH-----LIDKQCV 53
Qy 62 EKSEFDKYGGLADV--TFSIVYTN-----EFYEGRAAGASY--DAKAYOS 105
Db 54 PQTIDGYIDNM-DVHYTWSIPNSTNVKADPTMDFTLPQLALATDIAFVKSQGVGT 112
Qy 106 LTPGRVQA-----GTTD-----ANGVTVQLP-KKONG-----133

```

```

Db 113 ATVKATNQVTLVFSDDYVEKSDVKGELDFMTAFNQKITGNETVDLVPLENGTVIDV 172
Qy 134 ---KDAVTIKKEEPREGVVAATNMVAFV-----YEM 163
Db 173 EVGEKTPSPTELTFFYCGVNDASNPSLHVVVRVNVAKVNIIPAVFTDIIAGAKQTNFDS 232
Qy 164 IKQDSYKYGTEBELAVHVIYPRNVVA--NDG-----SLHVKVGTAEENG-----207
Db 233 IKAFTGYSADVFPTRNGAPISSSTNFSATSDGSVALGNLTDVQISYTTTDDGGSTQY 292
Qy 208 -----LNGAEFYISKS-----GSPGVKYIQGVKQGLYTWTTDKEQAKREFTGK 254
Db 293 DMTAKLAGTDFVTKQSTWTPASGGGEG-----GTTGSYTLTKEDAKTKATLEGAEF 346
Qy 255 EIGENDFTEAE-----NGGELTVKNLEVGSIYILEBKAPNNAELENTKPTPTTEANN 309
Db 347 KLVDSKGTVLQENITTNASGQLSIADLKEDTYQLLETKAPGYKL--DTPVEFTIGENN 404
Qy 310 QT-PVEKTVKNTSKVDKTTPTSLDGDVAIGEKIKQISVNIPLGIADKEGDKANKYKEN 368
Db 405 QAITTKENTLNTGSEVELTKLDAATKATLAGATPELQ-----DKEGNT--LQTD 451
Qy 369 LVDKADALTFPDNVTSGEYAVALYDQDT-----VIAPENYQVTEQANGFTVAV 416
Db 452 LKTDENGVLKATDVLPGSVQFVETSAPGTGKLDNSPVSFEVLAGETDQVVKTKENTLEY 511
Qy 417 NPAIYPTLTPG--GLTKFVYFMHLEKADPT-----KGRKN-----EAM 453
Db 512 GSEVELTKLDSATKATLAGATF--ELQDKEGNTLQTDLKTDENGVLKATDVLPGSVQFVETS 570
Qy 454 VDNGHDTDDQTPPTVEVNTS--GGRFVKVDGDVT-----ATQALAGASPVVD 498
Db 571 APTGKLDNSPVSFEVAGETDQVVKTKENTLEVGSVELTKLDSATKATLAGATPELQ 630
Qy 499 QNSDTANYLKIDETTKAATWVKTKAEATFTTTADGLVDITGLKYGTGYLEETVADDDYV 558
Db 631 KEGNTLQ-----TGLTIDENGVLKATDVLPGTVQFVETKADIGYE 670
Qy 559 LITNRIEFVNEQSYGTENLV-----SPEKVPNKHGTLPTSGKGIVYLGSGAVLLIAG 580
Db 671 IDTTPVSF---EIVAGEEDPIVKTKENTLVPTPVPPTPVPPTPVPPTPVP 727
Qy 581 -----SPEKVPNKHGTLPTSGKGIVYLGSGAVLLIAG 616
Db 728 YPTVPPTKPEVVPVVKTKENSEDESKTPIRITQSLPTGDTNSPAGLVIILALSLGS 787
Qy 617 VYFARR 622
Db 788 LLLKRR 793

```

## RESULT 9

F83901

Hypothetical protein BH2014 [imported] - *Bacillus halodurans* (strain C-125)  
 C.Species: *Bacillus halodurans*  
 C.Date: 01-Dec-2000 #sequence\_revision 01-Dec-2000 #text\_change 09-Jul-2004  
 C.Accession: F83901  
 R.Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hira  
 Nucleic Acids Res. 28, 4317-4331, 2000  
 A.Title: Complete genome sequence of the alkaliphilic bacterium *Bacillus halodurans* and  
 A.Reference number: AB9650; MUID:20512582; PMID:11058132  
 A.Accession: F83901  
 A.Status: preliminary  
 A.Molecule type: DNA  
 A.Residues: 1-1816 <STO>  
 A.Cross-references: UNIPROT:Q9KBB4; UNIPARC:UPI00000C3DBA; GB:AP001514; GB:BA000004; NIT  
 A.Experimental source: strain C-125  
 C.Gene: BH2014

```

Query Match 7.3%; Score 234.5; DB 2; Length 1816;
Best Local Similarity 20.2%; Pred. No. 8, 5e-05;

```

Matches 126; Conservative 80; Mismatches 221; Indels 197; Gaps 21;

```
QY 40 QLVIRHKKMTDLPPELLIONSGKMSFPDKYQGLADVTFISYVNTNFEYQRAAGASVDAA 99
Db 960 QSTVNTAKIVDPSP-----NQILLBDSFQ-----LFTSVNTAT----- 993
QY 100 KQAVOSLTPGKPVAAQT-----TDANGNTVTVQLPKQNGKDAVYITKEEPKGGVAAAT 152
Db 994 ---GTTKGAPLBERGVADYILEIFTEDEGNETFEL-SPPNNIDRAYIL--EYKSFYIAQN 1046
QY 153 NMVVAFPV-YEMIKOTDGSYKGTGTEELAVHVIYPRNVAND-GSLHVKVGTAEINEGANG 210
Db 1047 NEFVSNQVSPFGDELTMTESNTSSSIQVRLSGSGTSGRGSGIETVKVATSGETIAG 1106
QY 211 AEFVTSKSSGPPGVKKYIQGVKDGILYWT--DKEQAKRFITGSKSYEIGEND----- 260
Db 1107 ATFVLVDSTGEIVLRTVVTG-EDGIATPANLMYDDYILRETSAPEGYLWGINDESHITLN 1165
QY 261 -----FTEANGTGE-----L 271
Db 1166 TELHELVENAKIIRDVQLHKIDETGSLQGAFTLQQRQSDPFVDYRENLETDENGCV 1225
QY 272 TVKNLENGSYLIEEVKAPNNAELIENQTKTEFTIEANNQTEVEKTVKN-----DTSKV 324
Db 1226 AVEDLEPDQYQFEVAAPSGVVL--NADPIFSIGEQTGTEVILTKENEILLGSEVLKVV 1283
QY 325 DKTPSS--LDGKDAVIGEKIKYQISVNIPLGIADKEGAPKAYKFNLYDKDAALTPDNV 382
Db 1284 DRDDSSYLEGAE-----FSLDEGDNV--LQGLTTDDEGKLIVASDL 1324
QY 383 TSGEAYVALYDGDVTIAPENYQVTEQGANGFTVAVNPVAVIPTLTGCGTLKPFYFMHLNEKA 442
Db 1325 RPNYQPV-----ETTSFEHLDDTPYRFTIIVLDQEQQLVY----- 1362
QY 443 DPTKGFKNNAVNDNGHTDDQTPPEVENVYTGGRKFIKVDGVTATQALAGSFVVRDQNSD 502
Db 1363 -----ATNRLIPGSQVLKVDADHPAV-VLEGAFESFLDEDCN 1399
QY 503 TANYIAKIDETTKAATWVTKAEATFTTTAGLVNDITGLKXGYLLETVAAPDDVYLLTN 562
Db 1400 IVQ-----BGLSTDANGQLIVTDLRPNYQVETAPADEYIILDAT 1439
QY 563 RIEFVNEQSYGTTEMLVSPKVP 586
Db 1440 PREFTIERSQEGISIVENELIP 1463
```

RESULT 10  
AC1458  
surface anchored protein (LPXTG motif) [imported] - *Listeria innocua* (strain Clj11262)  
CISpecies: *Listeria innocua*  
CISDate: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 09-Jul-2004  
CISAccession: AC1458  
R:Glasier, P.; Frangoul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Biocker, J.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihl, H.; Jones, L.M.; Karet, U.  
Science 294, 849-852, 2001  
A:Authors: Krett, J.; Kuhn, M.; Kunst, F.; Kurapatk, G.; Madeno, E.; Maitournam, A.; Meak, C.; Schlueder, T.; Simoes, N.; Tietze, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlend, A.; Title: Comparative genomics of *Listeria species*.  
A:Reference number: AB1077; MUID: 21537279; PMID: 11679669  
A:Accession: AC1458  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-681 <GLA>  
A:Cross-references: UNIPROT:Q92PA4; UNIPARC:UPI00000CC163; GB:AL592022; PIDN:CA095435.1  
A:Experimental source: strain Clj11262  
CISGenetics:  
A:Gene: lin0202

**QY** 4 LKKVWT-VSTLLILPLF-TSVLGTTTAAEENGSAQLVHKKMTDLEDDPLIQNSG 61  
|||::: : :  
**Db** 8 LKRVLGLAFSLVLIATAVFQITIVKAATSYGSEBLNVE-----LIDDKGT 53  
::|::: : :  
**QY** 62 EKSEPFKYQGL-ADVFSIYNVTNEFEYGRRAAGASVDAAKQAVSLTGPKVAQGTTDAN 120  
::|::: : :  
**Db** 54 PKTDREYIMNMNVHYTWISIPNSTNV-----KAGDSMPALPS-QIALATDLAEFVKOSK 106  
::|::: : :  
**QY** 121 GNV-TVQLPKQNGDAVYT-----IKEEP-----KEGVVAATNMV-VAFPYDEM 163  
|||::: : :  
**Db** 107 GQTGGATATVKRATNQTVTFVSDEYKHSKDIKGEIDFWTTFNQKIYTGNEKINLEPIENS 166  
|||::: : :  
**QY** 164 -----IKOTDSYKYGTEE-----LAHVHIYPKNV----- 188  
|||::: : :  
**Db** 167 TTNVDNVGCKTPVSPLETIFKTGYWDANNPSLIHVVRYAQQNIENAVFTDIGAKQ 226  
|||::: : :  
**QY** 189 -----VAN-----DG-----SLHVKVGTAENE 206  
|||::: : :  
**Db** 227 TLNPDSIKAFHGTYSTRDVFTAGPISNTNFSKTSDGFSSVTGLNLDBSQSYTTTADIG 286  
|||::: : :  
**QY** 207 G-----LNGAEFVLSKF-----GSPTGVKKIQGVKGLTYWTTDKEQAK 246  
|||::: : :  
**Db** 287 GKSTOYDNTRAKLSGBDFVKQTSTWTPASGGGGAGDGT-----GSTVLTKODAKTK 338  
|||::: : :  
**QY** 247 REITGKSYE-----IGENDFTEANGTGELTVKMLEVGSYLIEEVKAPNNNELIENO 298  
|||::: : :  
**Db** 339 ALLEGAEFLVDANGTILOGENITTTDA---SGQLNVSNLKPATYOLLIETKAPTGYTL--DS 393  
|||::: : :  
**QY** 299 IYKPTPTI-EANNOTPEYKVANDTSKYDKTTPSLDGDVALIGEIKIKIOISVINPILGIADK 357  
|||::: : :  
**Db** 394 TEVEEFITGENNKKEIFTYTKENTLDTGAVELTKVDAAATAAIILAAG----- 435  
|||::: : :  
**QY** 358 BEDANKYVENLVKDKDALTFPDNVTSGEYAVALYDGDPTVIAPENYQVTEQANGFTVAVN 417  
|||::: : :  
**Db** 436 -----ATFELDOKGGSVLQDCLKTDENGILAKTVD---LVPGNYQ----- 471  
|||::: : :  
**QY** 418 PAYIPTLTBGGTLKFVYFMHLNFKADPTKGFKNEAVNDNGHTDDQPPTVEVV--TGGR 475  
|||::: : :  
**Db** 472 --FVEISAFTG-----YLDSSPVPFELIADB-----TDQIVKTKRNILEVGAVE 515  
|||::: : :  
**QY** 476 FIKVDGDVATQA-LAGASFVRVDQNSDTANYLIKIDETTKAATWTKTAABATPTTTDAG 534  
|||::: : :  
**Db** 516 LTKVD--AATKALLAGATFELQDKQKTLQ-----TGLTTDENG 552  
|||::: : :  
**QY** 535 LVDTIGLKGYTYLBEETVAPDDYVLLTNRIEFVNNEOSYGTENLVSEBKVPNKHKGTLP 594  
|||::: : :  
**Db** 553 VLKMTDLVPGTYQFVEVTNAPIGYELDASPVEF--EIYGENNQVVKVTK-ENKAETFP 608  
|||::: : :  
**QY** 595 ST 596  
|||::: : :  
**Db** 609 PT 610

RESULT 11  
G86643  
[hypothetical protein ybf (imported) - Lactococcus lactis subsp. lactis (strain IL1403)]  
CSpecies: Lactococcus lactis subsp. lactis  
CDate: 23-Mar-2001 #sequence\_rev10n 23-Mar-2001 #text\_change 09-Jul-2004  
CAccession: G86643  
RBolotin, A.; Wincker, P.; Manger, S.; Jalllon, O.; Malarme, K.; Weissenbach, J.; Ehrlich,  
Genome Res. 11, 733-753, 2001  
ATitle: The complete genome sequence of the lactic acid bacterium Lactococcus lactis ss  
AReference number: AB6625; MDID:2125186; PMID:1137471  
AStatus: preliminary  
AMolecule type: DNA  
AResidues: 1-1983 <STO>  
ACross-references: UNIPROT:09CUJ5; UNIPARC:UPI00000C676A; GB:AE005176; PID:g12723000; P  
AExperimental source: strain IL1403  
CIGenetics:  
AIGene: ybf

Query Match 7.1%; Score 228; DB 2; Length 1983;  
 Best Local Similarity 21.9%; Pred. No. 0.00022;  
 Matches 113; Conservative 63; Mismatches 152; Indels 186; Gaps 22;

107 TFGKP---VAQGTDTNAN-----VTVOL---PKQNGKDAVYTIKEEPKEG---VVA 150  
 1100 TGKDKDKLIREGDTDKNGDLKLGKGLKLVTVLLLETKAPG---YTLDELVLVKGKLTVA 1155  
 151 AINNVAVPVMEMIKOTDSYKKGTBELAVNHYPPNVAVNGDSLHVKKVYG-TAEN----- 205  
 1156 KDGVSNNPPIQVENE-----PKVI-----LKKVGLITENGKEV 1190  
 206 -GLNGAEFVISKESGSPGVTKYIGVKGGLYTTTDEKQAKRFITGKSYEIGENDPTA 264  
 1191 KNPLOGAEFKVLDSNQ-----VSGYEKLTSD----- 1218  
 265 ENGTEGLVYKNI.EVGSYILEEYKAPNNABLIENQYTPPTTEANNQTPVEKTVKNDTSKV 324  
 1219 --SSGNVTIEKLTPEKKSIVETKAP--AGYILDPTPIDFELKANEGLI----- 1263  
 325 DKTTBSDLGKQVAIGKIKYQISVNI-----PLGIADKESGANKVKNVLD 371  
 1264 -----PDINL--EKVYQSAQLIKHNSNGALSGLFKVIDKQNT--IQTNLS 1310  
 372 KHDAAITPDNWTSGEYAVLYDGDVYIAPENYQVTEQANGPTVAVNPAYIPTLTGGLTK 431  
 1311 DDDGKVTATGSLPGDVSFV-----ETKAPGTGILNTIPHFTISA----- 1350  
 432 FVYFHLNEKADPTKGFKNENAVNDGHTDDQTPFVAVTVGGRFIVDGDVATQALAG 491  
 1351 -----EEGQPOWVIASDNFIN-----YQGSABLIKQDNE--GGPLSG 1386

492 ASFVVRDQNSDPTANYIKIDETIKATWTKAETFTTTAGDVLDTGLKGTYYLEET 551  
 1387 AEFKIVD-----KAKTIOA-----NLTKGKQKVVVDGLAPGESFVET 1426

552 VAPDDVLLTNRIEFVNVESYGTTENLVSPKVPN 587  
 1427 KAPGTGIIINTNPVHFTISDKBQPTTIASDSFVN 1462

RESULT 12  
 A35400  
 surface protein T6 precursor (strain D471) - Streptococcus pyogenes  
 C:Species: Streptococcus pyogenes  
 C>Date: 31-Aug-1990 #sequence\_revision 31-Aug-1990 #text\_change 09-Jul-2004  
 C:Accession: A35400  
 R:Schneewind, O.; Jones, K.F.; Fischetti, V.A.  
 J. Bacteriol. 172, 3310-3317, 1990  
 A>Title: Sequence and structural characteristics of the trypsin-resistant T6 surface pro  
 A:Reference number: A35400; MUID:90264329; PMID:2188957  
 A:Accession: A35400  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-537 <SCH-  
 A:Cross-references: UNIPROT:P18401; UNIPARC:UPI00001368BD; GB:M32978; NID:g153842; EIDN:  
 C:Keywords: transmembrane protein

Query Match 6.8%; Score 220.5; DB 2; Length 537;  
 Best Local Similarity 22.4%; Pred. No. 8.4e-05;  
 Matches 128; Conservative 82; Mismatches 185; Indels 177; Gaps 31;

155 VVAFVYEMIKOTG-----SYKGTBELAVNHYPPNVAVNGDSLHVKKVGAENEGGLNG 210  
 40 VLYIKIGSGVNTNDSPFINFRY-ARGVSLTETGPTS-----QGITTIAN-GINT 87  
 211 AEFVISKESG--SPGVTKYI--QGVKGLYTTTDEKQAKRFITGKSYEIGENDPTA 264  
 86 GKIKPFTENVISNGTATYNNAGASVYIALTLGATDERTVPIILASV----- 137  
 265 ENGTEGLVYKNI.EVGSYILEEYKAPNNABLIENQYTPPTTEANNQTPVEK-TVKNDTSK 323  
 138 -NGEGLNLYTKNIDSKSNYL-----YQGTSAVAKSLPSITKK 172

324 VDKTTPSLDKQVAIGEKIKYQISVNIPLGIADKESGANK--YVKNELVJKHDAALTPDN 381  
 173 VTGIDTVNKKTTSLASVLSYSLTFELP--SYTKEA-VNKTIVVSDNM--SEGLTNNFS 227  
 382 VT-----SGEYAVLYDGDVYIAPENYQV--TEQANGPTVAVNPAYIPTLTGGLKVFVF 435  
 228 LTVEMKQKVAANTBDSVNV--ENTKIGIAKEVNNQFNLSFYDLSLSISPNISYKAV-- 283  
 436 MHLNEKA-----DFTKGFKNBAND-----NGHTD----- 460  
 284 --VNNKAIIVEEGNPNKAEFFYSNNPTKG-NTYDINDRKPDKNGKITSKEDSKIVTYTOY 340  
 461 -----DQTPFVAVTVGGRFIVDGDVATQALAGSFVVRQ 499  
 341 AFRKVDVSKTPLIGAFISGVYDTSNKLIDIVTNK-----NGYASTQVSSG-KYIKEL 394  
 500 NSDTANYIKIDETTKAATW---VTKA--EATTFPT---TADGLVDITGKGYTV--- 546  
 395 KAPKQSLNTEFYETIANNVATVTSANSKSTYTSDDKKAFTDNEQGWLNKGIIFYSI 454  
 547 -----YLBETVAPDDVLLTNRIEFVNVESYGTTENLVSPKVPNKKKGLTP 594  
 455 DSRPTGNDVKEAYIESTKA-----LTDGTFESKSESGT---VLEETDIPNTKLELP 505

595 SNGKGIYVY--LGSQAVLLIAGVYFARRK 624  
 506 STGSGITVLFKAIGSAMIGAI-GIYIVRRK 536

RESULT 13  
 A83695  
 hypothetical protein BH0361 [imported] - Bacillus halodurans (strain C-125)  
 C:Species: Bacillus halodurans  
 C>Date: 01-Dec-2000 #sequence\_revision 01-Dec-2000 #text\_change 09-Jul-2004  
 C:Accession: A83695  
 R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Maeni, N.; Fuji, F.; Hiri  
 Nucleic Acids Res. 28, 4317-4331, 2000  
 A>Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and  
 A:Reference number: A83650; MUID:20512582; PMID:11058132  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Accession: A83695  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-1661 <STO>  
 A:Cross-references: UNIPROT:Q9KFW0; UNIPARC:UPI00000C385A; GB:AP001508; GB:BA000004; NIT  
 A:Experimental source: strain C-125  
 C:Genetics:  
 A:Gene: BH0361

Query Match 6.7%; Score 216.5; DB 2; Length 1661;  
 Best Local Similarity 20.2%; Pred. No. 0.00074;  
 Matches 140; Conservative 74; Mismatches 225; Indels 253; Gaps 25;

62 EMSRFDKYQG--LADVTSIVVNTNEF-----YEQRAAGASV----- 96  
 1088 EYTKVDADRGVLAQTFTLVDSBGEFARTLETGEDKATFVNLLGYDLKEDSAPBG 1147  
 97 -----DAKQAVQSLTPEKPVAGCTTDAN-----GNVTVOLPKQNGK----- 134  
 1148 YLVGINDTORVYIDVTLHEVTENKSDINRVSAGVAGVQLQRYVDEBTGSLQALPALQO 1207  
 135 ---DAVYTI--KEEPKEGVAAATNNVAVPVMEMIKOTDSYKKYGBELAVY----- 181  
 1208 KYDDEFTVIAEMETBEGIVFAGSLB--PGYQFVELNAPVGYKIDETPPVFTVEEDRT 1264  
 182 -----HIYKVVAVNDGSLHVKKVGAENEG-LNGAEFVISKESGSPGVTKYIQGV 231  
 1265 EYTBLOKENVHLF-----GSVQLVKVDADDAANTLBSAEFTLLDGEQNV-----V 1309  
 232 KQGLYTTTDEKQAKRFITGKSYEIGENDPTAENGTEGLVYKNI.EVGSYILEEYKAPNN 291  
 1310 QEGLT---TTDE-----NGQVVVTDLKPGBYQFVETATAPAG 1341

```

OY      292 AELLINOTKPFTTEANNOTPEVKTIVKNDISKVDKTTPSLDGKVALGEKIKYIOISNIP 351
Db      1342 YEL--EAPPIGFTTERRNOOEVAIVAVENHLLPGSVOLVKVDAIAITLE-----GAB 1392
OY      352 LGIADKREGDANKYVKFENVVDKHAALTFDNTVTSGEYAAVLXDGPTVLAPENYQVTEQANG 411
Db      1393 FSLLDGEGNV---VQBGITPATGQVVVTDLKPEFYQV-----ETKAPAGIELDAPIG 1444
OY      412 FTVAVNPAYIPFTLLPGGTLKFVFMHLEKADPTKGPFKEANVDNGHTDDQPTPEVEVT 471
Db      1445 FTIERNOOEVAIVA-----VEN-----HILIP 1465
OY      472 GGGKRFIKDGVNTATQALAGASFVVRDQNSDTANYLKIDETTKAATVKTAEATITPTT 531
Db      1466 GSVOLIKDKRDSAV-TLEGAFSLDDEE---GNVLR-----EGIRTG 1504
OY      532 ADGVLDTIGLKKGTYLLEETVAPDDYVLLTNRIEFV-----568
Db      1505 ENGOLMVIDLKRGDIQFIETKAPGCIYELDDPIPIATIEKGTDEVVLTVFNKLTPEVPES 1564
OY      569 -----NEQSYGTENLVSP-----KVPNKHGK-----591
Db      1565 PENPENPENPENPEIPDELTEPESPENPETPESPXGHSDESEVPNSDKNPEKSDSNRRID 1624
OY      592 TLPSTGSGKITVYLGSGAVLLLIAGVYPAKR 623
Db      1625 RLPTQGEBEFLFLLILGLLIMTAGVLLLKKR 1656

```

## RESULT 14

hypothetical protein yhgE [imported] - Lactococcus lactis subsp. lactis (strain IL1403)  
C|Species: Lactococcus lactis subsp. lactis  
C|Date: 23-Mar-2001 #sequence\_revision 23-Mar-2001 #text\_change 09-Jul-2004  
C|Accession: F86719  
R|Botolin, A.; Winkler, P.; Manger, S.; Jallion, O.; Malarme, K.; Weissenbach, J.; Ehrlich  
Genome Res. 11, 731-753, 2001  
A|Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis s  
A|Reference number: AB6625; MUID:21235186; PMID:11337471  
A|Accession: F86719  
A|Status: preliminary  
A|Molecule type: DNA  
A|Residues: 1-614 <STG>  
A|Cross-references: UNIPROT:O9CHH3; UNIPARC:UPI00000DD43FC; GB:AE005176; P  
A|Experimental source: strain IL1403  
C|Genetics  
A|Gene: yhgE

| Query Match           | 6.7%             | Score 215;                                                     | DB 2;       | Length 614; |
|-----------------------|------------------|----------------------------------------------------------------|-------------|-------------|
| Best Local Similarity | 21.3%;           | Pred. No. 0.00021;                                             |             |             |
| Matches 155;          | Conservative 79; | Mismatches 269;                                                | Indels 226; | Gaps 38;    |
| QY                    | 1                | MKOLKKVYVYVSTLLILPLFTSVLGGTTAFAB-----ENGESAOVLVHKKKMDLP        | 52          |             |
| Db                    | 3                | INSLNKNPALASVSLALSTLAGPSGLATVNHKTTNVADQVSNGEVA---LYAQSAGNQ     | 59          |             |
| QY                    | 53               | DP-----LIONSGKEMSE-----FDKYQGLADYVF-----SIYVYVNE               | 85          |             |
| Db                    | 60               | NPTQNGGLVDBDGNVYNDSGIANNOGTGGAGNTSGTKNDISQOQLTBESTPSTMANVY--   | 117         |             |
| QY                    | 86               | FYEQRAGASVDAAKQAVQSLTPGKPVAGCTTDANGNVTVQPKKQNGKDAVYTIKKEEPK    | 145         |             |
| Db                    | 118              | FSATKYVGTGVPQT-VTDPSFTDITTSVAPVITDASG-----LADPTGLTGDGYLFHQ---  | 168         |             |
| QY                    | 146              | EGVAATITMVVAFPYFYEIMIKQTDGSYKGTETELAVVHIYPRKNVANDSLHVKKVGTLEN  | 205         |             |
| Db                    | 169              | ---VTVVNGITTT--VGDFIVQV---SHEDSQAGIVNYVPKLDMSSSAGLQTSATITNND-  | 218         |             |
| QY                    | 206              | EGLNG-----AEFVLSIKSGSPPTVYKIQGVQDGLVYWT--TIDKEQAKRFI           | 249         |             |
| Db                    | 219              | DNFNQGTPTQJLANPRAATGNSDQTLTNTDNNAGNENRLANG-----TWTNGSDNQQTTTAA | 272         |             |
| QY                    | 250              | TGKSYEIGEN--DPTLEANGTGLVTVNKLDEGVSYLEEVAPANNAAELINQTKPTTIE     | 306         |             |

```

Db      273 AGNTNMNVNTVFPSSQTNNGNGTTGV---TGYIYITD-QLPNN--LVNSSTVTVSTVI 325
QY      307 ANNOTPVEKTYKNDNTSKVDKTTPLSDGKVAIGEKIYOISVNIPLGIADKEGDANKYVK 366
Db      326 VN-----VTNMSGTKVGLTLPFTD-----YTLT-----NDGNGKIV 356
QY      367 FNLY--DKHDALTFDNTSGEYAYALYDGDYTIAPENIO--VTEQANGFTYAVNPAVI 421
Db      357 VTLLTTCGQOHAASLL-----GSADGAL--NIIIPSTVSAIGSATSDATYTIITNAY- 405
QY      422 PTLTPGGTLKVVYFMHLNEKADPTKGFKNENAVNGHHTDQTPPTVEVTVGSKRFIVDG 481
Db      406 ----GADL-----STTAVKSTLNV-----GGELEMTKDA 431
QY      482 DVTAQTOLAGASFPV--RDQNSDPANYLIKIDE-----TTKATV----- 519
Db      432 STNA--ALAGTFTFVVRADNKEDNODPEVANAIAFNNSASGITYTNLTSKAAAVTGDTS 489
QY      520 ---KTKAEA--TFTTADGLVDITGLKY-----GTYLEETVAPDDYVL----- 559
Db      490 GNATTSATAPFTFTTGKDGIAITFNGILNVDNNTDGSNTVVYLVLEVAAPGCVQLPSVTA 549
QY      560 --LNNRIEFVNEOSYGTTENLVSPKVPKHKGTLPSGTGKGIYVLGSGAVLLIAGV 617
Db      550 ANLGAVALT-----ASTAPAADTTITNNKPPALPFTGQGLAGIILAIATVSGVIA-- 599
QY      618 YFARRKEN 626
Db      600 FAIRKRKN 608

```

## RESULT 15

A:Internal protein, probable peptidoglycan bound protein (LPTXG motif) homolog lin2495  
 C:Species: *Listeria innocua*  
 C:Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 09-Jul-2004  
 C:Accession: AB1744  
 R:Classer, P.; Frangoul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Biochecker,  
 ; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fshhi, H  
 D.; Jones, L.M.; Karet, U.  
 Science 294, 849-852, 2001  
 A:Authors: Krett, J.; Kuhn, M.; Kunst, F.; Kurapak, G.; Madheno, E.; Maitournam, A.; M  
 ok, C.; Schluener, T.; Simoes, N.; Tierrey, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlend  
 A:Title: Comparative genomics of *Listeria* species.  
 A:Reference number: AB1077; MUID:21537279; PMID:11679669  
 A:Accession: AB1744  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-940 <GUA>  
 A:Cross-references: UNIPROT:O928N7; UNIPARC:UPI00000CCGB3; GB:AL592022; PIDDN:CA097722.1  
 A:Experimental source: strain Clp11262  
 C:Genetics:  
 A:Gene: lin2495

|    | Query Match           | 6.6%                                                         | Score 212.5       | DB 2       | Length 940 |  |
|----|-----------------------|--------------------------------------------------------------|-------------------|------------|------------|--|
|    | Best Local Similarity | 23.5%                                                        | Fred. No. 0.00053 |            |            |  |
|    | Matches 160           | Conservative 77                                              | Mismatches 256    | Indels 189 | Gaps 37    |  |
| QY | 18                    | LPLFTSVLGTITTAFAENGESAOLVYHKKKQWTDLPDIIQNSGKEMSEFDRYQGLADVTF | 77                |            |            |  |
| DB | 370                   | VPLSISIVSQPIKSEKTNISAVNVFYEDENGDELASETISGK---SGENYQ--TTKRTI  | 425               |            |            |  |
| QY | 78                    | SIYVNTNEFYQRPAG--ASVDAKQAQVOSTLPKPKPAQGTITDANGNVTVQLPKKQNGK- | 134               |            |            |  |
| DB | 426                   | TNYLJKE--IEGQASGGQFGSDSTVTVVYBEKADGAPVTVKVVDDGN-ELATSDTLNGKI | 482               |            |            |  |
| QY | 135                   | DAVY-----TIKEP--KSGVVAATMMVVAFPVYEMIKQTDSDSKYGTGTEBLAVNH     | 182               |            |            |  |
| DB | 483                   | DAPVOSTAKSTITGMVAKQTTTPANANGVTFNNAQITVY-VYEKADGAPAVVKKVYDED  | 536               |            |            |  |
| QY | 183                   | TPKRVVANDGSLHKV-----KVGTAENGL-----NGAEFVYSKSEG               | 220               |            |            |  |



```
Db 537 ---GNEIATSDTLNGKIDAPYOSTAKSITGMVAVKTPANANGVFTNANQTVTVYVEKADG 593
QY 221 SPGTIKYIGVNDGLYTTTDEQAK-----RFTGKSYEIGENDPTEANGTSEL 271
Db 594 APVTYKYVD--BDGNEIATSDTLNGKIDAPYOSTAKSITGMVTK-----TTPANANGVF 645
QY 272 TVKNLEVGSYILEEVK-APNNAELIR---NOTKPTIIBANNQTPVEKTVKNDTSKYDKT 327
Db 646 TNDNQTV-TVYVEKASGAPVTYKYVDEBDGNEIATPDTLNGKIDAPYOSTAKSLSGMAVKI 704
QY 328 TPSLDGKDVAISEKIKYQISVNIPLGIADKEG--DANKYKFNLVKXHDALTFDNTS 384
Db 705 TP-----ANATGVFTDNTQTVY-VYERVDGAPVTYKYVD 738
QY 365 GEYAYALVNDGTVIAPEN-----YOTEOA-NGFTVANVPAYIPTL--TPGGLKEV 433
Db 739 G-----DGNELATPDTLNGKIDAPYOSTAKSITGMVTKTTPANANGVFTDNTQTVTVY 791
QY 434 YFMHINEKADPTKGFKNBANVDNGHTDQTPPTVEVVTGAKRFIKVGDVTAQALAGAS 493
Db 792 Y-----EKA-----DGAPVTYKYVDG-----DGNELAT----- 814
QY 494 FVVRDQNSDTPANYLKDITTKAAT-----W--VKTKAATFTTTADGIVDITGLKYGT 545
Db 815 -----PDTLNG-KLD-TSYAATAKNLSGMKLTATPANATGVFTTDAQTVTVVYAKQED 865
QY 546 YLEETVAPDDYVLLTNRIEFVNNEOSYGTENTLVSPEKVPKNKHGTLPSGKG----I 601
Db 866 DPKEDKTPSPNTQPDCKTTIKINENKPNPNTSK---FTTIKKQTK--LPKTDGNOQESIL 919
QY 602 VYILGSGAVLLIAGVYFARR 623
Db 920 FGLIGTCFVLL--GIYSVSKK 938
```

Search completed: July 6, 2006, 02:37:23  
Job time : 47 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.9  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: July 6, 2006, 02:28:16 ; Search time 301 Seconds

(without alignments)  
1926.860 Million cell updates/sec

Title: US-10-661-809a-13

Perfect score: 3225

Sequence: 1 MKQLKKVWTVSTLLILPL.....GAVLLIAGVFPARRKXENA 627

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2849598 seqs, 925015592 residues

Total number of hits satisfying chosen parameters: 2849598

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Maximum Match 0%

Listing first 45 summaries

Database : UniProt 7.2.\*

1: uniprot\_sprot.\*

2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description   |
|------------|-------|-------------|--------|----|---------------|
| 1          | 3202  | 99.3        | 625    | 2  | 083618_ENTRPA |
| 2          | 2446  | 75.8        | 625    | 2  | 03YD3_ENTFEC  |
| 3          | 458   | 14.2        | 507    | 2  | 048707_LACIE  |
| 4          | 438.5 | 13.6        | 476    | 2  | 083619_ENTFEC |
| 5          | 423   | 13.1        | 658    | 2  | 03YD4_ENTFEC  |
| 6          | 395   | 12.2        | 473    | 2  | 03YD2_ENTFEC  |
| 7          | 355.5 | 11.0        | 565    | 2  | 03YD9_ENTFEC  |
| 8          | 355   | 11.0        | 554    | 2  | 03DE5_STRAG   |
| 9          | 355   | 11.0        | 554    | 2  | 03K250_STRAG  |
| 10         | 352   | 10.9        | 554    | 2  | 03DB92_STRAG  |
| 11         | 352   | 10.9        | 554    | 2  | 08B089_STRAG  |
| 12         | 352   | 10.9        | 554    | 2  | 08B089_STRAG  |
| 13         | 345   | 10.7        | 693    | 2  | 03YD2_ENTFEC  |
| 14         | 337.5 | 10.5        | 674    | 2  | 03D871_STRAG  |
| 15         | 334   | 10.4        | 502    | 2  | 03D871_STRAG  |
| 16         | 334   | 10.4        | 502    | 2  | 03K0A5_STRAG  |
| 17         | 334   | 10.4        | 502    | 2  | 08A441_STRAG  |
| 18         | 333   | 10.3        | 705    | 2  | 03D92_STRAG   |
| 19         | 333   | 10.3        | 705    | 2  | 08DVR6_STRAG  |
| 20         | 329.5 | 10.2        | 674    | 2  | 08B4C4_STRAG  |
| 21         | 323   | 10.0        | 525    | 2  | 08B4C4_STRAG  |
| 22         | 296.5 | 9.2         | 665    | 2  | 08B4C4_STRAG  |
| 23         | 296   | 9.2         | 1530   | 2  | 04E006_LISMO  |
| 24         | 296   | 9.2         | 1530   | 2  | 08Y479_LISMO  |
| 25         | 292   | 9.1         | 564    | 2  | 072X94_BACCI  |
| 26         | 292   | 9.1         | 564    | 2  | 03DMP5_STRAG  |
| 27         | 282.5 | 8.8         | 554    | 2  | 04MW04_BACCI  |
| 28         | 279.5 | 8.7         | 553    | 2  | 081D71_BACCI  |
| 29         | 277   | 8.6         | 323    | 2  | 03D870_STRAG  |
| 30         | 276.5 | 8.6         | 454    | 2  | 02PB82_STRAG  |
| 31         | 276.5 | 8.6         | 490    | 2  | 06NKO5_CORDI  |

|    |       |     |      |   |              |                    |
|----|-------|-----|------|---|--------------|--------------------|
| 32 | 272   | 8.4 | 553  | 2 | 04MU04_BACCI | 04MU04 bacillus ce |
| 33 | 267   | 8.3 | 1599 | 2 | 04SEW1_LISMO | 04SEW1 listeria mo |
| 34 | 267   | 8.3 | 1512 | 2 | 08Y591_LISMO | 08Y591 listeria mo |
| 35 | 266   | 8.2 | 1102 | 2 | 06H0B3_BACCI | 06H0B3 bacillus an |
| 36 | 266   | 8.2 | 3333 | 2 | 06H0B7_BACCI | 06H0B7 bacillus th |
| 37 | 261   | 8.1 | 1697 | 2 | 071X17_LISMO | 071X17 listeria mo |
| 38 | 260   | 8.1 | 1697 | 2 | 04ELD4_LISMO | 04ELD4 listeria mo |
| 39 | 259.5 | 8.0 | 3393 | 2 | 0630P2_BACCI | 0630P2 bacillus ce |
| 40 | 255.5 | 7.9 | 1163 | 2 | 04MMH3_BACCI | 04MMH3 bacillus ce |
| 41 | 255   | 7.9 | 546  | 2 | 02Z2P8_STRAG | 02Z2P8 streptococ  |
| 42 | 252.5 | 7.8 | 1508 | 2 | 073CV3_BACCI | 073CV3 bacillus ce |
| 43 | 252.5 | 7.8 | 1975 | 2 | 06SN14_BACCI | 06SN14 bacillus th |
| 44 | 252   | 7.8 | 3232 | 2 | 03F0J5_BACCI | 03F0J5 bacillus th |
| 45 | 251.5 | 7.8 | 901  | 2 | 03DMP6_STRAG | 03DMP6 streptococ  |

## ALIGNMENTS

|                           |                                                                         |                           |
|---------------------------|-------------------------------------------------------------------------|---------------------------|
| RESULT 1                  | 083618_ENTRPA                                                           | PRELIMINARY; PRT; 625 AA. |
| ID                        | 083618_ENTRPA                                                           |                           |
| AC                        | 083618                                                                  |                           |
| DT                        | 01-JUN-2003, integrated into UniProtKB/TrEMBL.                          |                           |
| DT                        | 01-JUN-2003, sequence version 1.                                        |                           |
| DT                        | 07-FEB-2006, entry version 12.                                          |                           |
| DE                        | Cell wall surface anchor family protein.                                |                           |
| GN                        | OrderedLocNames=EF1093; ORFName=EF_1093;                                |                           |
| OS                        | Enterococcus faecalis (Streptococcus faecalis).                         |                           |
| OC                        | Bacteria; Firmicutes; Lactobacillales; Enterococcaceae; Enterococcus.   |                           |
| OX                        | NCBI_TaxId=1351;                                                        |                           |
| RN                        | [1]                                                                     |                           |
| RP                        | NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].                          |                           |
| RC                        | STRAIN=Y583 / ATCC 700802;                                              |                           |
| RX                        | MEDLINE=22550857; PubMed=12663927; DOI=10.1126/science.1080613;         |                           |
| RA                        | Paulsen I.T., Banerjee L., Myers G.S.A., Nelson K.E., Seshadri R.,      |                           |
| RA                        | Read T.D., Fouts D.R., Eisen J.A., Gill S.R., Heidelberg J.F.,          |                           |
| RA                        | Tetzelin H., Dodson R.J., Umayam L.A., Brinkac L.M., Beanan M.J.,       |                           |
| RA                        | Daugherty S.C., DeBoy R.T., Durkin S.A., Kolonay J.F., Madupu R.,       |                           |
| RA                        | Nelson W.C., Vamathevan J.J., Tran B., Upton J., Hansen T., Shetty J.,  |                           |
| RA                        | Khouri H.M., Uettermann T.R., Radune D., Ketchum K.A., Dougherty B.A.,  |                           |
| RA                        | Frazer C.M.,                                                            |                           |
| RT                        | "Role of mobile DNA in the evolution of vancomycin-resistant            |                           |
| RT                        | Enterococcus faecalis."                                                 |                           |
| RL                        | Science 299:2011-2074(2003).                                            |                           |
| CC                        | Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms |                           |
| CC                        | Distributed under the Creative Commons Attribution-NonDerivs license    |                           |
| CC                        | EMBL; AB016830; AAO80893.1; -; Genomic DNA.                             |                           |
| DR                        | TIGR; EF1093;                                                           |                           |
| DR                        | BioCyc; EFAB226185:EF1093-MONOMER; -                                    |                           |
| DR                        | GO; GO:0009986; C:cell surface; IEA.                                    |                           |
| DR                        | InterPro; IPR008454; Cna B.                                             |                           |
| DR                        | InterPro; IPR001899; Gram_pos_anchor.                                   |                           |
| DR                        | Pfam; PF05738; Cna B; 3.                                                |                           |
| DR                        | Pfam; PF00746; Gram_pos_anchor; 1.                                      |                           |
| DR                        | TIGRfam; TIGR01167; LfXTG_anchor; 1.                                    |                           |
| DR                        | PROSITE; PS50847; GRAM_POS_ANCHORING; 1.                                |                           |
| KW                        | Cell wall; Complete proteome; Peptidoglycan-anchor.                     |                           |
| SQ                        | SEQUENCE 625 AA; 67989 MW; EB97450B111F0DE CRC64;                       |                           |
| Query Match               | 99.3%; Score 3202; DB 2; Length 625;                                    |                           |
| Best Local Similarity     | 99.7%; Pred. No. 1,66-161;                                              |                           |
| Matches 625; Conservative | 0; Mismatches 0; Indels 2; Gaps 1;                                      |                           |
| QY                        | 1 MKQLKKVWTVSTLLILPLFTSVLGTTPAFABENGSAQVYHKKMTDLPPLIONSG 60             |                           |
| DB                        | 1 MKQLKKVWTVSTLLILPLFTSVLGTTPAFABENGSAQVYHKKMTDLPPLIONSG 60             |                           |
| QY                        | 61 KEMSEFDKYGSLADYVFSTYNTNEFYRORAAAGAYDAKAVOSITPKPVAOGTTDAN 120         |                           |
| DB                        | 61 KEMSEFDKYGSLADYVFSTYNTNEFYRORAAAGAYDAKAVOSITPKPVAOGTTDAN 120         |                           |

QY 121 GNVTVQLPKKONGKAVYTIKEEPKRGVVAATNMVAPVPMIKOTDGSYKGTIELAV 180  
DB 121 GNVTVQLPKKONGKAVYTIKEEPKRGVVAATNMVAPVPMIKOTDGSYKGTIELAV 180  
QY 181 VHIYPRNVVANDGSLHVKKVGTAENGLNGAEFVLSKSGSPGVKYIGVVDGLYTWTT 240  
DB 181 VHIYPRNVVANDGSLHVKKVGTAENGLNGAEFVLSKSGSPGVKYIGVVDGLYTWTT 238  
QY 241 DKEQAKRPTGKSYEIGENDPTEAENGTELVKLVKLEVSYLEEVKAPNNMELLENOK 300  
DB 241 DKEQAKRPTGKSYEIGENDPTEAENGTELVKLVKLEVSYLEEVKAPNNMELLENOK 298  
QY 301 TPEFTLEANNQTPVEKTVKNDTSKVDKTPSLDGDVAIEKIKYOISVNIPLGIADKEG 360  
DB 299 TPEFTLEANNQTPVEKTVKNDTSKVDKTPSLDGDVAIEKIKYOISVNIPLGIADKEG 358  
QY 361 ANKRYKFNVLVDKHDALTFDNTVSGEYAYALYDGTVAIPENVQVTEQANGFTVAVNPAY 420  
DB 359 ANKRYKFNVLVDKHDALTFDNTVSGEYAYALYDGTVAIPENVQVTEQANGFTVAVNPAY 418  
QY 421 IPTLPFGTGLKRVYPMHLEKADPTKGFQNEANVNGHDDQTPPVVYVVGKRPVKYD 480  
DB 419 IPTLPFGTGLKRVYPMHLEKADPTKGFQNEANVNGHDDQTPPVVYVVGKRPVKYD 478  
QY 481 GGVATQALAGASPVVRDQNSDPTANYLKIDETTKATVYKTAEATFTTTADGLVDITG 540  
DB 479 GGVATQALAGASPVVRDQNSDPTANYLKIDETTKATVYKTAEATFTTTADGLVDITG 538  
QY 541 LKGYTYLEETVAPDDYVLLTNRIEFVNVESGYTTEINVSPKVPNKHKGLPTSGKG 600  
DB 539 LKGYTYLEETVAPDDYVLLTNRIEFVNVESGYTTEINVSPKVPNKHKGLPTSGKG 598  
QY 601 IYVYLGSGAVLLIAGVYFARRKENA 627  
DB 599 IYVYLGSGAVLLIAGVYFARRKENA 625

RESULT 2  
Q3Y3D3\_ENTFC PRELIMINARY; PRT; 625 AA.  
AC Q3Y3D3;  
DT 11-OCT-2005, integrated into UniProtKB/TrEMBL.  
DT 11-OCT-2005, sequence version 1.  
DT 07-FEB-2006, entry version 4.  
DE Surface protein from Gram-positive cocci, anchor region precursor.  
GN ORFNames=EfAeBAPF\_2569;  
OS Enterococcus faecium DO.  
OC Bacteria; Firmicutes; Lactobacillales; Enterococcaceae; Enterococcus.  
OX NCBI\_TaxID=333849;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=DO;  
RG US DOE Joint Genome Institute (JGI-PGF);  
RA Copeland A., Lucas S., Lapidus A., Barry K., Dettler C., Glavina T.,  
RT Hammon N., Israeli S., Pitluck S., Richardson P.;  
RT "Sequencing of the draft genome and assembly of Enterococcus faecium  
DO.";  
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=DO;  
RG US DOE Joint Genome Institute (JGI-ORNL);  
RA Larimer F., Land M.;  
RT "Annotation of the draft genome assembly of Enterococcus faecium DO.";  
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=DO;  
RG US DOE Joint Genome Institute (JGI-PGF);  
RA Copeland A., Lucas S., Lapidus A., Barry K., Dettler C., Glavina T.,  
RT Hammon N., Israeli S., Pitluck S., Richardson P.;  
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.

CC -! CAUTION: The sequence shown here is derived from an  
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is  
CC preliminary data.  
CC -----  
CC Distributed under the UniProt Consortium, see <http://www.uniprot.org/terms>  
CC Copyrighted by the Creative Commons Attribution-NonDerivs license  
CC -----  
CC EMBL: AAK0300002; EAN11034.1; -? Genomic\_DNA.  
DR GO:0009986; C:cell surface; IEA.  
DR GO:0005618; C:cell wall; IEA.  
DR InterPro: IPR008454; Cna\_B.  
DR InterPro: IPR001899; Gram\_pos\_anchor.  
DR Pfam: PF00746; Gram\_pos\_anchor; 1.  
DR TIGRPFAM: TIGR01167; LpXTG\_anchor; 1.  
DR PROSITE: PS00847; GRAM\_POS\_ANCHORING; 1.  
KW Cell wall; Peptidoglycan-anchor; Signal.  
FT SIGNAL 1 33 Potential.  
SQ SEQUENCE 625 AA; 68172 MW; D4653B1755F241A2 CRC64;  
Query Match 75.8%; Score 2446; DB 2; Length 625;  
Best Local Similarity 74.2%; Pred. No. 2,1e-121;  
Matches 462; Conservative 64; Mismatches 95; Indels 2; Gaps 1;  
QY 1 MKQLKKVYVSTLLIPLFTSVLGTTTAFANEENGESQVLVHKKKMTDLPDPLIONSG 60  
DB 1 MKGHKKINMVLGVLLPLPLTNSFGAKKVFAEET-AAQVILHKKKMTDLPDPLIONSG 58  
QY 61 KEMSEFDKIQGLADYTPSYNTNFEYQRAAGASVDAKQAVQSLTGTGKPAQGTTPAN 120  
DB 59 KEMSEFDYQGLADISFSYVNTQEFVARDGASVDAKQAVQSLTGTGKPAQGTTPAN 118  
QY 121 GNVTVQLPKKONGKAVYTIKEEPKRGVVAATNMVAPVPMIKOTDGSYKGTIELAV 180  
DB 119 GNVTVQLPKKONGKAVYTIKEEPKRGVVAATNMVAPVPMIKOTDGSYKGTIELAV 178  
QY 181 VHIYPRNVVANDGSLHVKKVGTAENGLNGAEFVLSKSGSPGVKYIGVVDGLYTWTT 240  
DB 179 IHLVPMNTVGNQDGLTKTKTGTAENBALNGAEFVLSKSGSPGVKYIGVVDGLYTWTT 238  
QY 241 DKEQAKRPTGKSYEIGENDPTEAENGTELVKLVKLEVSYLEEVKAPNNMELLENOK 300  
DB 239 DQTKAKHFTTGHSHYDIGNDFAEASIEKGQLLVNHLVGVKXLEEVKAPNNMELLENOK 298  
QY 301 TPEFTLEANNQTPVEKTVKNDTSKVDKTPSLDGDVAIEKIKYOISVNIPLGIADKEG 360  
DB 299 TPEFTLEANNQTPVEKTVKNDTSKVDKTPSLDGDVAIEKIKYOISVNIPLGIADKEG 358  
QY 361 ANKRYKFNVLVDKHDALTFDNTVSGEYAYALYDGTVAIPENVQVTEQANGFTVAVNPAY 420  
DB 359 QNKYTFKLIDTHDALTFDNDSSGTVAVALYDGNKEIDPVNVSVTEQDGTFSVDPMY 418  
QY 421 IPTLPFGTGLKRVYPMHLEKADPTKGFQNEANVNGHDDQTPPVVYVVGKRPVKYD 480  
DB 419 IPTLPFGTGLKRVYPMHLEKADPTKGFQNEANVNGHDDQTPPVVYVVGKRPVKYD 478  
QY 481 GGVATQALAGASPVVRDQNSDPTANYLKIDETTKATVYKTAEATFTTTADGLVDITG 540  
DB 479 GGVATQALAGASPVVRDQNSDPTANYLKIDETTKATVYKTAEATFTTTADGLVDITG 538  
QY 541 LKGYTYLEETVAPDDYVLLTNRIEFVNVESGYTTEINVSPKVPNKHKGLPTSGKG 600  
DB 539 LKGYTYLEETVAPDDYVLLTNRIEFVNVESGYTTEINVSPKVPNKHKGLPTSGKG 598  
QY 601 IYVYLGSGAVLLIAGVYFARRK 623  
DB 599 IYVYLGSGAVLLIAGVYFARRK 621

RESULT 3  
Q48707\_LACIE PRELIMINARY; PRT; 507 AA.  
ID Q48707\_LACIE  
AC Q48707;



```

QY 297 NOTKPTLEANNQTEVEKTVKNDTSKVDKTTPSLDGKQVAIGEKIKQISVINPIGLIAD 356
DB 188 -----PPEKTVTVKDKQGGFQ-----GEMINQVLTQIIPANILG 221
QY 357 KEGDANKYKFNLDVKHDAALTFEDNVTSGEYVALYDGDVTJAPENYQVTEQANGFTYAV 416
DB 222 -----YQEFRLSDKADTTLTL-----LPESIEVKVAKGVTTT--GYTLTQKHGFTLDF 268
QY 417 NPAYIPITLPGSTLKFVYFMHLEKADPRTKFKMEAN--VDNGHTDDQTPPTVEVTVGGR 475
DB 269 SIKDQUNRA-NQTMVTSYQMRLEKTAEPDTAINNEGQVLTDRHT--LTKRAVTRGKGS 324
QY 476 FIKVDGVATATQALAGASFVVDSDPTANVLIKIDETTKAATWTKKAEATFTTTADGL 535
DB 325 FVKVDSE-NAKTLLEAVFIVKNGQGEYIN-----ETANGYMQKEKALAKKFTSNQGE 378
QY 536 VDTGLKGTYYLLEETVADDDVLLTNRIEFPVNEQSYGTENLVSPKVEN--KHKGT 592
DB 379 FSVKGLKDGQYFLBEISAPKGYLLNQTEIPFTVGKNSYATNGQRTAPFLHIVKKVSGSF 438
QY 593 LPSTG-GKGIYVYLGSAVLLLIAGV-----YFARBRKE 625
DB 439 LPKTNBRSIMULTIAG---LILIGVVIWLFYQKQKGE 474

```

## RESULT 5

```

Q3Y0Y4_ENTFC PRELIMINARY; PRT; 658 AA.
ID Q3Y0Y4_ENTFC PRELIMINARY; PRT; 658 AA.
AC Q3Y0Y4;
DT 11-OCT-2005, integrated into UniProtKB/TrEMBL.
DT 11-OCT-2005, sequence version 1.
DE 07-FEB-2006, entry version 3.
DE Surface protein from Gram-positive cocci, anchor region precursor.
GN ORFNames=BfaeDRaft_1904;
OS Enterococcus faecium DO.
OC Bacteria; Firmicutes; Lactobacillales; Enterococcaceae; Enterococcus.
OX NCBI_TaxID=333849;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=DO;
RG US DOE Joint Genome Institute (JGI-ORNL);
RA Larimer F., Land M.;
RT "Annotation of the draft genome assembly of Enterococcus faecium DO.";
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=DO;
RG US DOE Joint Genome Institute (JGI-ORNL);
RA Larimer F., Land M.;
RT "Annotation of the draft genome assembly of Enterococcus faecium DO.";
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=DO;
RG US DOE Joint Genome Institute (JGI-PGF);
RA Copeland A., Lucas S., Lapidus A., Barry K., Dettler C., Glavina T.,
  Hamon N., Israni S., Pritch S., Richardson P.;
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
CC -! CAUTION: The sequence shown here is derived from an
  EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
  preliminary data.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivs license
CC -----
CC EMBL: AAK0300014; EMBL0156.1; -! Genomic DNA.
CC GO: GO:000986; C:cell surface; IBA.
CC InterPro: IPR008454; Cna_B.
CC InterPro: IPR001899; Gram_pos_anchor.
CC Pfam: PF05738; Cna_B_2.
CC TIGRfams: TIGR01167; LPXTG_anchor; 1.

```

```

KW Signal.
FT SIGNAL.
SQ SEQUENCE 658 AA; 71457 MW; F540948280207BAA CRC64;
Query Match 13.1%; Score 423; DB 2; Length 658;
Best Local Similarity 25.7%; Pred. No. 4,7e-14;
Matches 188; Conservative 101; Mismatches 254; Indels 188; Gaps 38;

```

```

QY 2 KQKKVWTVSTLLILFLPSVLTGTTAFABE-----NGESAQVVIHKKKMTDLPDP- 55
DB 5 RYKKAAGMLIGMLGIACGPGQY---RAYADVTQKTPPEKVNITVH-KLMYDQGTOLN 59
QY 56 -----IQNSGKMESEF---DKYQ---GLADYFYSYNTNEYF---EGRPAASVDAKQ 101
DB 60 VDIDGIRKNDGYTHDPYPPGVTKYNNKADGDVEFTLGNITDQVLPEDSDLTAKYDEIK 119
QY 102 AVQSILPQGPVAAQGT-----DANGVT-VQLPKONGKDAVTTIK-EPEGVVA-- 150
DB 120 DVEDGSGNSEYVKANTKITSVDNGEITFPADQPAVYNSKGNVYVESKSAAGLVQK 179
QY 151 ATNNVVAFPVYEMIKQDGSYKYGTEELAVVHIYRKNVAVANGSLHKV--GTAENE- 206
DB 180 AKPMVVIAPM-----TDNT---GSGFLKDHLHYKNIYSK-LSFELTKFGDDGTAQSKQ 229
QY 207 -GLNGAEFVTSKESGSPGVKXKIQGVDGLYTTTDEKQAKRFYKSYEIGENDTEAE 265
DB 230 TPLKGAKEFLYKGE-----PGKGTGLGD-----LYS 255
QY 266 NOTGELTVKNTLEVGSIYILBEYKAPNNAELIENQTTPT-----IEANQTPVEK 315
DB 256 DDQGLKLTATDITLTKYFEVEVP---SEVVGSDKEPFPADQYLLGADARNDAHNLTEI 311
QY 316 TVKNDTSK-----VDKTPSLDGKD--VAIGEKIKYQISVINPIGLIADKEGDA 361
DB 312 TNDGVSILKASVYNYKAPVLDKTYTNGTGQSHSQIGDAVYQGTIHPDIA--GGA 368
QY 362 NKYKFNLDVKHDAALTFEDNVTSGEYVALYD-GDT-----VI 398
DB 369 D-----GITVNGVSETPSPYVFPMGDTAQSLSYAKANIKVTNKGSGV 415
QY 399 APEN--YQVTEQANGFTV--AVNPAYIPTLP--GGLKRFYFMHLEKADPTGFKNE 451
DB 416 LKENIDYKIQNSENFVIDFIVNNGQVSDYVASLHGDLQMTNMYVNDVSAVANPLTNS 475
QY 452 AN-VDNGHTDDO---TPPTVEVTVGKRFIKVDGDVAT-----QALAGASFVVDSD 502
DB 476 VDFVTNNPFPNQEHEHETKADVTVYGAFLKVDSGLFGTGIKATPLGSAEPAK--NAE 533
QY 503 TANYLKIDET---TKAATWTKKAEATFTTTADGLVDITGLKYGYLYLEETVADDDV 558
DB 534 GKYYGGLVDTDKDGKCAVWVDVANAAIILKSDKEGHFEITGLTEGEYSLEETKAPENYQ 593
QY 559 LITNRIEFPVNAEQSYGTENLVSPKVENYKHKGTLPSTGKGIYY-----LQSGAVL 611
DB 594 KLTKEISFVKDQSY-KEENRIT---IKNNQASVPMTGSGNGFQTYVILISCLLLAGA-- 647
QY 612 LLIAGVFAAR 622
DB 648 -LSAVVYFKKK 657

```

## RESULT 6

```

Q3Y3D2_ENTFC PRELIMINARY; PRT; 473 AA.
ID Q3Y3D2_ENTFC PRELIMINARY; PRT; 473 AA.
AC Q3Y3D2;
DT 11-OCT-2005, integrated into UniProtKB/TrEMBL.
DT 11-OCT-2005, sequence version 1.
DE 07-FEB-2006, entry version 3.
DE Surface protein from Gram-positive cocci, anchor region:Cna B-type
  precursor.
GN ORFNames=BfaeDRaft_2570;
OS Enterococcus faecium DO.
OC Bacteria; Firmicutes; Lactobacillales; Enterococcaceae; Enterococcus.

```



Best Local Similarity 22.7%; Pred. No. 1.5e-10; Matches 159; Conservative 98; Mismatches 221; Indels 223; Gaps 30;

QY 5 KKVWTVSTLLILPLFSLVLTGTTAFPAENGESAO-----LVHKKMTDLPDLION 58  
DB 3 KKMIVTGSVLLLVTAGIGIFGNGKLVRKA--EVAQKPEEVITLHKKGFSSVPEER-PN 59  
QY 59 SKKEMSEPKYQGLDVPFSINVTNEFE-----QRAAGASVDAKQAVOS----- 105  
DB 60 SLVSTIDGEE-ENIEGVDPDLPDVEVYDILRDNPLTEREDGLASAIIEWIQKRTTE 118  
QY 106 -----LTPKPRVAQGTIDANGV---TVQLPKK-QNGKDAVYTIKE--EPKEGVAAAT 152  
DB 119 SWFLKRLTL---SIDKQTNEAGAVFSIVQYTEAPSRDKVYLLELFYSPAHISRLAS 175  
QY 153 NMVAVAPFYEMIKQIDGSY-----KYTEELAVHHIYKPKVNVANDGSLHKYKGTABN 205  
DB 176 PAVVMPVPM-MEDMVDVGDGSTMKDYNTD---VHLVYKN----- 212  
QY 206 EGLNGAEFVTSKESGPGTVKVIYIGVKDGLYMTTDDKEQAKRPITGSKYEIGENDPTEAE 265  
DB 213 -----EIRBAD----- 218  
QY 266 NGTGLVTKNLEVGSYILEEVKAPNNABLIENQTKPTTEANNQTPVEKTYKNDTSKYD 325  
DB 219 -----KQNNVESDLRQVTIIN-----EAGEQETISYI----- 246  
QY 326 KTTPLSDGDVAIGEKIKQISVINPLGIADKEGDANKYK-PNLVDKDAALLTPDNVTS 384  
DB 247 -----DLERKASTYITAPIPYFIDSLVLENGSAVINKYKIDTPTVGLTY----- 292  
QY 385 GEVAVAYLD-----GDTVIAP-ENVQYTEOANGPTAV-----NPAVIPTL-----TP 426  
DB 293 -----YDQELVRAGETIITLKGDYIYEVVNSNGVVTIILTBENAVAVDLGLRADAR 345  
QY 427 GGTLEFVYFMHLEKADPTKGFKNBANVDNGHTD-----DQTPPEVVTGSKRFLKVDG 481  
DB 346 GGDLTITVYLKVTLEADDFHNTFAVIEIGRNDPEDEGEVPEPKVTTGGRKFEKY-- 403  
QY 482 DVTATQALAGASFVVRDQNSDTANY-----LKIDETKRAATVTKAEITTTTAA 532  
DB 404 DASSELMLDARFEL--WNEDRSEVAIFYKGSPLAVYESGADRLEWATSGATEFVADG 461  
QY 533 DCLVITGKXGYTYYLEETVAPDDYVLLTNR---IEFVNVEOSYGTTEMLV-----SPK 584  
DB 462 NCYFEVQGLDYGTQMKETMAPEGVLLPTGEAAFTFIISYGSYNELIQVGVBNGPBR 521  
QY 585 VNNKKGTLPTSGKGIYVYLGSAVLLLIAGVYFARRKE 625  
DB 522 VNNMKRGLSPATGNGGLAFLILIGISLMI--GAYSWYRKSK 560

RESULT 8  
Q3DFE5\_STRAG PRELIMINARY; PRT; 554 AA.  
AC Q3DFE5;  
DT 22-NOV-2005, integrated into UniProtKB/TrEMBL.  
DT 22-NOV-2005, sequence version 1.  
DE 07-FEB-2006, entry version 4.  
DE Cell wall surface anchor family protein.  
GN ORFNames=SAM\_0661;  
OS Streptococcus agalactiae C3B11.  
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
OC Streptococcus  
NCBI\_TaxID=342617;  
RN NCB1  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=C3B11;  
RX PubMed=16172379; DOI=10.1073/pnas.0506758102;  
RA Tetteelin H., Masignani V., Cieslewicz M.J., Donati C., Medini D.,  
Ward N.L., Anguinoi S.V., Crabtree J., Jones A.L., Durkin A.S.,  
DeBoy R.T., Davidson T.M., Mora M., Scaletti M., Margarit y Ros I.,  
Peterson J.D., Hausen C.R., Sundaram J.P., Nelson W.C., Madupu R.,

RA Brinkac L.M., Dodson R.J., Rosevitz M.J., Sullivan S.A.,  
RA Daugherty S.C., Haft D.H., Selengut J., Gwinn M.L., Zhou L., Zafar N.,  
RA Khouri H., Radune D., Dmitrov G., Watkins K., O'Connor K.J.,  
RA Smith S., Usterback T.R., White O., Rubens C.B., Grandi G.,  
RA Madoff L.C., Kasper D.L., Telford J.L., Wessels M.R., Rappuoli R.,  
RA Frazer C.M.;  
RT Genome analysis of multiple pathogenic isolates of Streptococcus  
RT agalactiae: implications for the microbial 'pan-genome'.  
RL Proc. Natl. Acad. Sci. U.S.A. 102:13950-13955(2005).  
CC -!- CAUTION: The sequence shown here is derived from an  
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is  
CC preliminary data.  
CC  
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>  
CC Distributed under the Creative Commons Attribution-NonDerivs License  
CC  
CC EMBL; AAJ00100012; EAO73672.1; -; Genomic DNA.  
DR GO; GO:0009986; C:cell surface; IEA.  
SQ SEQUENCE 554 AA; 60150 MW; 64498CB80A075ED1 CRC64;

Query Match 11.0%; Score 355; DB 2; Length 554;  
Best Local Similarity 27.3%; Pred. No. 1.5e-10; Matches 164; Conservative 67; Mismatches 225; Indels 144; Gaps 29;

QY 125 VOLPKKQNGKDAVYTI-----KEPKEGVAAATNMVAVAPVE--MIKQIDGSKYKTEEL 178  
DB 1 MKLSKKLLFSAVAVLTIVAGSTVEPAQFATGMSIYRAAEVSEGRPAKTTVINIYKLQADS- 59  
QY 179 AVVHHIYKPNVNVANDGSLHKYKVTAEENGNGAEFVTSKESGPGTVKVIYIGVKDGLYTW 238  
DB 60 -----YKSEIITSNGGI-----ENKQGE--VISNYAKLGDNVGKLGQVQPKRYKV 101  
QY 229 TTD--KEQAKRIT-----GKSYEIGENDPTEANGTGLVTKNLEVGSYI---LE 284  
DB 102 KTDISVDELKTLTVEADAKVGTILBEGVS-LPKTKAQG-LVVDALDSSKNVAYLYVE 159  
QY 285 EVK-APNN-----AEL-IEHQTKPTTEANNQTPVEKTYKNDTSKYDKTTPSLD 332  
DB 160 DLKNSPNSITKAYAVPFLVLELVANSTGTFLEIN---IYKQNVVTDEPKTDKQVKGL- 215  
QY 333 GKQVA---IGEKIKQISVINPLGIADKEGDANKYKRNLDVKDAALTFDNV----- 382  
DB 216 GQDDAGYITGEFFKFKFLKSTIPANLGD-----TEKEEITDKFADGLYKSVGKIKIGS 268  
QY 383 -TSGEYAVAYLDGIV-----IAPENY-QVTEQANGPTAVANPAVIPTL-----PG 427  
DB 269 KTLRNDERTIDEPVDNONTIKITFKPEKFEIABLKGMTLVKQDALDQATNTDQA 328  
QY 428 GTLKEFVYFMHLEKADPTKGFKNBANVDNGHTDQ-----PPTVEVVTGSKRFLKV 479  
DB 329 AFLBIPVASTINEKAVLGKAIENITPELDYDTPDKADNPKPSPNPKPEVHTGGRFVVK 388  
QY 480 DGDVTAATQALAGASFVVRDQNSDTANYLKIDETTKAATVTKAATY----- 527  
DB 389 DS--TETQTLGAEF-----DLASDQ--AVKWTDALIKANTNKYIAGEAVTG 434  
QY 528 -----FTTADAGLVDTGLKYG-----TYYLEETVAPDDYVLLTNRIEFVNVEOSY 573  
DB 435 QPIKLSHTDGTFFELKGLAYVANDABGTAIVYLLKETKAPBGVYIIPKEIEIFYQSISY 494  
QY 574 GT--TENLV-----SPKVPNKGKGLPTSPSGKGIYVYLGSAVLLLIAGVYFARRKEN 626  
DB 495 NTKPFDITVDSADAPPTIKNNKRSRISPTGTGIGTAIFVALGAAVMAFVKMKERTKDN 554

RESULT 9  
Q3K250\_STRAI PRELIMINARY; PRT; 554 AA.  
AC Q3K250;  
DT 08-NOV-2005, integrated into UniProtKB/TrEMBL.  
DT 08-NOV-2005, sequence version 1.  
DE 07-FEB-2006, entry version 4.  
DE Cell wall surface anchor family protein.



Ordered locus names=SAK\_0776; ORF names=SAK\_0776;  
OS Streptococcus agalactiae serotype Ia.  
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
OC Streptococcus.  
NCBI\_TaxID=355315;  
[1]  
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].  
STRAIN=909 / ATCC 27591 / Serotype Ia;  
PubMed=16172379; DOI=10.1073/pnas.0506758102;  
RX PubMed=16172379; DOI=10.1073/pnas.0506758102;  
RA Tectelin H., Maignani V., Cieslewicz M.J., Donati C., Medini D.,  
Ward N.L., Anguioi S.V., Crabtree J., Jones A.L., Durkin A.S.,  
DeBoy R.T., Davidson T.M., Mora M., Scarselli M., Margarit Y Ros I.,  
Peterson J.D., Hauser C.R., Sundaram J.P., Nelson W.C., Madupu R.,  
Brinkac L.M., Dodson R.J., Rosovitz M.J., Sullivan S.A.,  
Daugherty S.C., Haft D.H., Selengut J., Gwinn M.L., Zhou L., Zafar N.,  
Khouli H., Radune D., Dimitrov G., Watkins K., O'Connor K.J.,  
Smith S., Ueberback T.R., White O., Rubens C.B., Grandi G.,  
Madoff L.C., Kasper D.L., Telford J.L., Wessels M.R., Rappuoli R.,  
Frazer C.M.;  
RA "Genome analysis of multiple pathogenic isolates of Streptococcus  
agalactiae: implications for the microbial 'pan-genome'";  
Proc. Natl. Acad. Sci. U.S.A. 102:13950-13955(2005).  
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>  
CC Distributed under the Creative Commons Attribution-NonDerivs license  
-----  
EMBL: CP000114; ABA44713.1; -; Genomic DNA.  
DR TIGR: SAK\_0776; -;  
DR GO: GO:0009986; C:cell surface; IEA.  
DR InterPro: IPR008454; Cna\_B.  
DR InterPro: IPR001899; Gram\_pos\_anchor.  
DR Pfam: PF00738; Cna\_B; 1.  
DR Pfam: PF00746; Gram\_pos\_anchor; 1.  
DR TIGR: TIGR01167; LpxTG\_anchor; 1.  
KW Complete proteome.  
SQ SEQUENCE 554 AA; 60150 MW; 64498CB80A075BD1 CRC64;  
  
Query Match 11.0%; Score 355; DB 2; Length 554;  
Best Local Similarity 27.3%; Pred. No. 1.5e-10;  
Matches 164; Conservative 67; Mismatches 225; Indels 144; Gaps 29;  
  
125 VOLPKKQNGDAVYTI-----KEEPKEGVAAATMVAAPVYR--MIKQDGSYKYTEEL 178  
1 MKLSKLLFSAVLTIVAGSTVBPVQAQATGMSIVAAAEVSGRPAAKTIVNIYKQADS- 59  
179 AVVHIYPRKVVANDGSLHVKVGTAEENGANGAEFVYSKESGPGTVKXIQQYKQGLYTM 238  
60 -----YKSEITSNGGI-----ENKDG--VISNYAKLGDVVKGLQGVQFRKYK 101  
  
239 TTD--KEQAKRFT-----GKSYEIGENDPTEANGTGLTVKNLEVGSI-----LE 284  
102 KTDISVDELKLLTVEADAKVGTILEEGVS-LPQRTNAG-LVVDALDSKSVRYLYVE 159  
285 EVK-APNN-----AEL-IEHOTKPTPTIEANNQTPVEKTVKNDTSKYDKTTPSID 332  
160 DLKNSNSNTTKAAYAPVLELPPANSTGGLSEIN---IYKNNVTDEPKTDKQKXKL- 215  
333 GKDAV---IGEKIKQIYSNIPGLIADKEGDKANKYKFNLVKHDALTFDNY----- 382  
216 GODDAGYTTIGBEFKFPLKSTIPANLGD-----YEKFEITDKFADGLTYKSVGKIKIGS 268  
383 -TSGEVAVLALYGDY-----IAPENY-OVTEQANGFTVAVPAIPILT-----PG 427  
269 KTLNDEHTTIDEPYDNONTLKITFKPEKFEIALLKMTILVKNQDADLAKATANTDA 328  
428 GTLKFPYFPAHLEKADPTGKFKNEANVNDGHTDQ-----PTVEVVTGSKKFIKY 479  
329 APLEIVASTINKAVLGAIENTFELQYDHTPDKADNPESNPRAKPEVHTGKKFVKK 388  
480 DGDVTAITQALAGASFVVRDONSPTANYLKIDETTKAATVVKAKFT----- 527  
389 DS--TETQTLGGAEF-----DLASDGT--AVKWTDALIKANTKNVYIAGEAVYTG 434

QY 528 ----FTTADGLVDITGLKYG-----TYYLEETVAPDDYVLLTNRIEFVNEQSY 573  
DB 435 QPKLKSHTDGFREIIGLAVYADANAEGTAVYTKAKETAPBEVYLPDKEIETFTVQTSY 494  
QY 574 GT--TENLY-----SPEKVPNKRKGTLPSTGKGIYVYLGSAVLLIAGVFAARRKEN 626  
DB 495 NTFPTDITVDASADATPDITKNNRPSIPNTGGIGTAFVAILGAANVAFVAKGMRRTKDN 554  
  
RESULT 10  
Q3DB92 STRAG PRELIMINARY; PRT; 554 AA.  
ID Q3DB92 STRAG  
AC Q3DB92;  
DT 22-NOV-2005. Integrated into UniProtKB/TrEMBL.  
DT 22-NOV-2005. sequence version 1.  
DT 07-FEB-2006. entry version 4.  
DE Cell wall surface anchor family protein.  
GN ORFNames=SAN\_0698;  
OS Streptococcus agalactiae COH1.  
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
OC Streptococcus.  
NCBI\_TaxID=342616;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=COH1.  
RX PubMed=16172379; DOI=10.1073/pnas.0506758102;  
RA Tectelin H., Maignani V., Cieslewicz M.J., Donati C., Medini D.,  
Ward N.L., Anguioi S.V., Crabtree J., Jones A.L., Durkin A.S.,  
DeBoy R.T., Davidson T.M., Mora M., Scarselli M., Margarit Y Ros I.,  
Peterson J.D., Hauser C.R., Sundaram J.P., Nelson W.C., Madupu R.,  
Brinkac L.M., Dodson R.J., Rosovitz M.J., Sullivan S.A.,  
Daugherty S.C., Haft D.H., Selengut J., Gwinn M.L., Zhou L., Zafar N.,  
Khouli H., Radune D., Dimitrov G., Watkins K., O'Connor K.J.,  
Smith S., Ueberback T.R., White O., Rubens C.B., Grandi G.,  
Madoff L.C., Kasper D.L., Telford J.L., Wessels M.R., Rappuoli R.,  
Frazer C.M.;  
RA "Genome analysis of multiple pathogenic isolates of Streptococcus  
agalactiae: implications for the microbial 'pan-genome'";  
Proc. Natl. Acad. Sci. U.S.A. 102:13950-13955(2005).  
RT -!- CAUTION: The sequence shown here is derived from an  
EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is  
preliminary data.  
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>  
CC Distributed under the Creative Commons Attribution-NonDerivs license  
-----  
EMBL: AAJ0100002; BA076719.1; -; Genomic DNA.  
DR GO: GO:0009986; C:cell surface; IEA.  
SQ SEQUENCE 554 AA; 60168 MW; 64498D4B49729AD1 CRC64;  
  
Query Match 10.9%; Score 352; DB 2; Length 554;  
Best Local Similarity 27.2%; Pred. No. 2.2e-10;  
Matches 163; Conservative 68; Mismatches 225; Indels 144; Gaps 29;  
  
125 VOLPKKQNGDAVYTI-----KEEPKEGVAAATMVAAPVYR--MIKQDGSYKYTEEL 178  
1 MKLSKLLFSAVLTIVAGSTVBPVQAQATGMSIVAAAEVSGRPAAKTIVNIYKQADS- 59  
179 AVVHIYPRKVVANDGSLHVKVGTAEENGANGAEFVYSKESGPGTVKXIQQYKQGLYTM 238  
60 -----YKSEITSNGGI-----ENKDG--VISNYAKLGDVVKGLQGVQFRKYK 101  
  
239 TTD--KEQAKRFT-----GKSYEIGENDPTEANGTGLTVKNLEVGSI-----LE 284  
102 KTDISVDELKLLTVEADAKVGTILEEGVS-LPQRTNAG-LVVDALDSKSVRYLYVE 159  
285 EVK-APNN-----AEL-IEHOTKPTPTIEANNQTPVEKTVKNDTSKYDKTTPSID 332  
160 DLKNSNSNTTKAAYAPVLELPPANSTGGLSEIN---IYKNNVTDEPKTDKQKXKL- 215  
333 GKDAV---IGEKIKQIYSNIPGLIADKEGDKANKYKFNLVKHDALTFDNY----- 382  
216 GODDAGYTTIGBEFKFPLKSTIPANLGD-----YEKFEITDKFADGLTYKSVGKIKIGS 268  
383 -TSGEVAVLALYGDY-----IAPENY-OVTEQANGFTVAVPAIPILT-----PG 427  
269 KTLNDEHTTIDEPYDNONTLKITFKPEKFEIALLKMTILVKNQDADLAKATANTDA 328  
428 GTLKFPYFPAHLEKADPTGKFKNEANVNDGHTDQ-----PTVEVVTGSKKFIKY 479  
329 APLEIVASTINKAVLGAIENTFELQYDHTPDKADNPESNPRAKPEVHTGKKFVKK 388  
480 DGDVTAITQALAGASFVVRDONSPTANYLKIDETTKAATVVKAKFT----- 527  
389 DS--TETQTLGGAEF-----DLASDGT--AVKWTDALIKANTKNVYIAGEAVYTG 434





Db 650 QRIENK-KVTIPOTGGTIGTILFTIIGLSIMLGAVIMKRROSEEA 693

RESULT 14

Q3DERS STRAG PRELIMINARY; PRT; 674 AA.

AC Q3DERS;

DT 22-NOV-2005, integrated into UniProtKB/TrEMBL.

DT 22-NOV-2005, sequence version 1.

DT 07-FEB-2006, entry version 5.

DE Cell wall surface anchor family protein.

GN ORFNames=SAM\_1372;

OS Streptococcus agalactiae CUB111.

OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;

OC Streptococcus.

OC NCBI\_TaxID=342617;

OX

RN

RP NUCLEOTIDE SEQUENCE.

RC STRAIN=CUB111;

RX PubMed=16172379; DOI=10.1073/pnas.0506758102;

RA Tetteijn H., Maignani V., Cielesiewicz M.J., Donati C., Medini D., Ward N.L., Angiuoli S.V., Crabtree J., Jones A.L., Durkin A.S., DeBoy R.T., Davidson T.M., Mora M., Scarselli M., Margarit y Ros I., Peterson J.D., Hauser C.R., Sundaram J.P., Nelson W.C., Madupu R., Brinkac L.M., Dodson R.J., Rosovitz M.J., Sullivan S.A., Daugherty S.C., Haft D.H., Selengut J., Gwinn M.L., Zhou L., Zafar N., Khouri H., Radune D., Dimitrov G., Watkins K., O'Connor K.J., Smith S., Uterback T.R., White O., Rubens C.E., Grandi G., Madoff L.C., Kasper D.L., Telford J.L., Wessels M.R., Rappuoli R., Frazer C.M.;

RT "Genome analysis of multiple pathogenic isolates of Streptococcus agalactiae: implications for the microbial 'pan-genome'.";

RL Proc. Natl. Acad. Sci. U.S.A. 102:13950-13955(2005).

CC -!- CAUTION: The sequence shown here is derived from an EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is preliminary data.

CC

CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>

CC Distributed under the Creative Commons Attribution-NoDerivs License

CC

DR EMBL, AAJ001000018; EAO73402.1; -; Genomic\_DNA.

DR GO; GO:0009986; C:cell surface; IEA.

DR GO; GO:0005618; C:cell wall; IEA.

DR KW Cell wall; Peptidoglycan-anchor.

SQ SEQUENCE 674 AA; 73217 MW; 5463FC096BCF0022 CRC64;

Query Match 10.5%; Score 337.5; DB 2; Length 674;

Best Local Similarity 25.1%; Pred. No. 1.7e-09;

Matches 192; Conservative 78; Mismatches 267; Indels 227; Gaps 39;

QY 1 MKQLKVVVTSSTLLILPLFTSVLGTTPAPAEENGESQVLVHKKKMTDLDPDLIN-- 58

DB 1 MKKIKKCLTMTSTLLIL--TSLPSVAPAFADD-ATTVTYLHKIV--PQAAFDNFT 53

QY 59 -----SGKEMSEFDKYQGLADVT-----FSIYVNT-NEFEYGRAGASVDA-- 98

DB 54 EGTGKNDSDVYKQINDLKSYFGSTDAKEIKGAFVFNERTGKRTEN--GKEVDLLE 111

QY 99 AKQAVQSLTPGKPVACQTTDANGN--TVQLPKQNGKDAVTTIIE--PKEGVAAATN 153

DB 112 AKDA-----EGGAVVLSGLTKDNGFVFNATK-----KGIVQIVELKEKSNVYDNNGLSDS 162

QY 154 MVVAPPVYEMIKOTGOSYKYGTBEELAVVHIYPKNVAN--DGLVHVKKVGTAEENGANG 210

DB 163 KAVPVKITPLVYNNQGVK-----DAHIYFRNTETKQVDKRNPKDKLDITDNRKDG 215

QY 211 AEFVSKSESGSPGVKYIYGK-----DGLVTWTTDKQAKRFTIGKSYEL- 256

DB 216 ---VVSATVGD--KKEYIVGTKLKSGDYKGLVMTDSMTKGLTPNNNVKVLTDGEDFVTL 270

QY 257 -----GENDPTEAENGSELTV-----KNLEVSQYIIIEYKAPRNALIE-NQTKTF 303

DB 271 NYKLVTDGQFRLATNATGLAAVAAAADKDVDEIKITYSATVNGSTTVEIPETNDVKLDY 330

QY 304 TIEANNOT-----PVEKTVKNDTSKYVDKTTPLSDKQVAIGEKIKYQISVNIPLGIADKEG 359

DB 331 ---GNPTEBSEFPQGTFRANQEIYI-----KDMVNDGIT----- 363

QY 360 DANKYK--FNLVDKHAALTPDNVTSGEYAYALYDGDVTAPEN---YQVTEQANGFTV 414

DB 364 DANVAVKAIPTLQEKQTDG--TWNVVASHATKPSRFEHTFTGLDNAKTVRVERVSGYT- 421

QY 415 AVNPAYIPLTPPGTLKTVYFMHLEKADPTKGFQNEANV--DNHHTDQ-----PPTVEV 469

DB 422 ---PEYV-----SFGNGVVTIKNNKNSNDPPTINSEPKV 453

QY 470 VTGGRFPIKVDGYATQALAGASFVRDQNSDTN-----YLKIDETT 513

DB 454 VTYGRKFPYKT--NQANTERLAGATTLVKKEGKYLARKGAATAEKAANKTKAKLADDEV 511

QY 514 KA-----AT-----WVTKA--EATPTTTAD 533

DB 512 KAVNDLTKEQEGQGTALATVDQKQAYNDAPVKANYSEWVADKKADNVKLIISNAG 571

QY 534 GLVDTTGKLYGTYLIEETVAPDDYLLNRIEFPVNBOSY--GTENLV-----SPE 583

DB 572 GQFELTGLDKGTYGLEEQAPAGVATLSGDVFEVTATSYSKATTDIAYDKSVKDAQ 631

QY 584 KVPNGHKGLPSTGKGIYVYVLSGAVLLILAGVYFARRRENA 627

DB 632 QYQK-KVTIPOTGGTIGTILFTIIGLSIMLGAVIMKRROSEEA 674

RESULT 15

Q3DRT1 STRAG PRELIMINARY; PRT; 502 AA.

AC Q3DRT1;

DT 22-NOV-2005, integrated into UniProtKB/TrEMBL.

DT 22-NOV-2005, sequence version 1.

DT 07-FEB-2006, entry version 4.

DE Cell wall surface anchor family protein, putative.

GN ORFNames=SAN\_1518;

OS Streptococcus agalactiae COH1.

OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;

OC Streptococcus.

OX NCBI\_TaxID=342616;

RN

RP NUCLEOTIDE SEQUENCE.

RC STRAIN=COH1;

RX PubMed=16172379; DOI=10.1073/pnas.0506758102;

RA Tetteijn H., Maignani V., Cielesiewicz M.J., Donati C., Medini D., Ward N.L., Angiuoli S.V., Crabtree J., Jones A.L., Durkin A.S., DeBoy R.T., Davidson T.M., Mora M., Scarselli M., Margarit y Ros I., Peterson J.D., Hauser C.R., Sundaram J.P., Nelson W.C., Madupu R., Brinkac L.M., Dodson R.J., Rosovitz M.J., Sullivan S.A., Daugherty S.C., Haft D.H., Selengut J., Gwinn M.L., Zhou L., Zafar N., Khouri H., Radune D., Dimitrov G., Watkins K., O'Connor K.J., Smith S., Uterback T.R., White O., Rubens C.E., Grandi G., Madoff L.C., Kasper D.L., Telford J.L., Wessels M.R., Rappuoli R., Frazer C.M.;

RT "Genome analysis of multiple pathogenic isolates of Streptococcus agalactiae: implications for the microbial 'pan-genome'.";

RL Proc. Natl. Acad. Sci. U.S.A. 102:13950-13955(2005).

CC -!- CAUTION: The sequence shown here is derived from an EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is preliminary data.

CC

CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>

CC Distributed under the Creative Commons Attribution-NoDerivs License

CC

DR EMBL, AAJ001000022; EAO75904.1; -; Genomic\_DNA.

DR GO; GO:0009986; C:cell surface; IEA.

SQ SEQUENCE 502 AA; 53450 MW; E3B7F5603F6D6192 CRC64;

Query Match 10.4%; Score 334; DB 2; Length 502;

Best Local Similarity 26.3%; Pred. No. 1.7e-09;

Matches 150; Conservative 79; Mismatches 223; Indels 118; Gaps 26;

```
QY 100 KQAVQSL-----TRGKVAQCTTANGNVTVQLPKQNGKDAVTTIKKEPKGV 148
Db 3 KKNIQSLVSLAFGNAVSPVPIAPAA--ETGITVQDTPQKATYKAYRFPDAEIDNAN 60
QY 149 VAATNNVVAFPVYEMIKQTDGSKYKGTBELAVVHIYPKNVANDGSLHVKVGTAH-NEG 207
Db 61 VSDSNDDGA--SLIIPQGEAEYKASTDFNSLP-----TTTNGGRYVTKQDTASANE- 112
QY 208 LNGAEFVISKSESSPGTVKXIQGVKQGLYTWTTDKQAKRFITGKSYEIGENDFTEAEN- 266
Db 113 -----IATWA-----KSI SANTPV--STYTESNND 136
QY 267 GTGELTVNKNLEVGSYLLEBVKAPNNNAELINQTKTPTTIANNOPTVEKTVKNDTSKVDK 326
Db 137 GTEVIVNS--QYGYVYVS--STVNGAVIMVTSVTP-----NATIEK--NTDATWGDG 184
QY 327 TTPSLDGKQVAIGEKIKYQISVNIPLGIADKEGDANKYKFNLVDKHDAALTFDNYTSGE 386
Db 185 GCKTVQKTYSVGDTYKTYTITTKAAYNHYGTE---KVYQYIKDTMPSASVVD-LNEGS. 239
QY 387 YAVVALYDSDTVI-----APENYQVTEQANGFTVAVNPAYIPTLTPOG----- 428
Db 240 YEVITIDGSGNITTLTQSEKATGKYNLENNNFITTI--PWAATNTPGTONTONGANDD 297
QY 429 -----TLKFVYFEMHLEKADPTKG----FKYEAENVDNQHTDDQTPTYEVVYTGKRFI 477
Db 298 FFYKGINITITVYTVGLKSGAKPGSADLPENTNIATINPNTSNDPQOKTVARDGOITIK 357
QY 478 KYDGDVATQALAGASFVVRDQNSDTANYLKIDETAKATWVKTKAATFTTTADQLVD 537
Db 358 KIDGSTKA--SIQGAIFVLKNA---TGQFLNENDTNH-VEW--GTEANATEYTTGADGIIT 410
QY 538 ITGLKXGTYLLEBTVAPDDYVLLTNRIEFVYNQSYGTT--ENLVSPKVPNKHGKT-LP 594
Db 411 ITGLKEGTYLVLEKKAPLGYNLLDNSQKVLIGDGTDTTNSDNLVNPTEVN-NKGTELP 469
QY 595 STGGKGIYVYVLSGANVLLIAGVYFARRK 624
Db 470 STGGIGTTIFYIIGALVIGAGIVLVARRR 499
```

Search completed: July 6, 2006, 02:36:34  
Job time : 306 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.9  
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: July 6, 2006, 02:48:51 ; Search time 166 Seconds  
(without alignments)  
1561.462 Million cell updates/sec

Title: US-10-661-809a-13

Perfect score: 3225  
Sequence: 1 MKQLKKWYTVSTLLIPL.....GAVLLINGVPRARRKENA 627

Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2097797 seqs, 463214858 residues

Total number of hits satisfying chosen parameters: 2097797

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA Main:

1: /EMC\_Celerra\_SIDS3/ptodata/2/pubpaa/US07\_PUBCOMB.pep:\*  
2: /EMC\_Celerra\_SIDS3/ptodata/2/pubpaa/US08\_PUBCOMB.pep:\*  
3: /EMC\_Celerra\_SIDS3/ptodata/2/pubpaa/US09\_PUBCOMB.pep:\*  
4: /EMC\_Celerra\_SIDS3/ptodata/2/pubpaa/US10A\_PUBCOMB.pep:\*  
5: /EMC\_Celerra\_SIDS3/ptodata/2/pubpaa/US10B\_PUBCOMB.pep:\*  
6: /EMC\_Celerra\_SIDS3/ptodata/2/pubpaa/US11\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description                            |
|------------|-------|-------------|--------|-------|----------------------------------------|
| 1          | 3225  | 100.0       | 627    | 3     | US-09-071-035-218 Sequence 218, App    |
| 2          | 3225  | 100.0       | 627    | 4     | US-10-206-576-218 Sequence 218, App    |
| 3          | 3225  | 100.0       | 627    | 5     | US-10-661-809-13 Sequence 13, Appl     |
| 4          | 3225  | 100.0       | 627    | 5     | US-10-912-362-218 Sequence 218, App    |
| 5          | 2892  | 89.7        | 560    | 3     | US-09-071-035-220 Sequence 220, App    |
| 6          | 2892  | 89.7        | 560    | 4     | US-10-206-576-220 Sequence 220, App    |
| 7          | 2892  | 89.7        | 560    | 5     | US-10-912-362-220 Sequence 220, App    |
| 8          | 2446  | 75.8        | 625    | 4     | US-10-661-809-19 Sequence 19, Appl     |
| 9          | 438.5 | 13.6        | 476    | 4     | US-10-661-809-11 Sequence 11, Appl     |
| 10         | 395   | 12.2        | 473    | 4     | US-10-661-809-17 Sequence 17, Appl     |
| 11         | 347   | 10.8        | 430    | 3     | US-09-071-035-274 Sequence 274, App    |
| 12         | 347   | 10.8        | 430    | 4     | US-10-206-576-274 Sequence 274, App    |
| 13         | 347   | 10.8        | 430    | 5     | US-10-912-362-274 Sequence 274, App    |
| 14         | 334   | 10.4        | 502    | 4     | US-10-333-002-16 Sequence 16, Appl     |
| 15         | 311   | 9.6         | 384    | 3     | US-09-071-035-276 Sequence 276, App    |
| 16         | 311   | 9.6         | 384    | 4     | US-10-206-576-276 Sequence 276, App    |
| 17         | 311   | 9.6         | 384    | 5     | US-10-912-362-276 Sequence 276, App    |
| 18         | 296.5 | 9.2         | 665    | 3     | US-09-769-787-127 Sequence 127, App    |
| 19         | 296.5 | 9.2         | 665    | 5     | US-10-472-928-778 Sequence 778, App    |
| 20         | 296.5 | 9.2         | 665    | 5     | US-10-873-528-127 Sequence 127, App    |
| 21         | 296   | 9.2         | 1530   | 6     | US-11-045-004-34 Sequence 34, Appl     |
| 22         | 267   | 8.3         | 1612   | 6     | US-11-045-004-38 Sequence 38, Appl     |
| 23         | 266   | 8.2         | 1788   | 4     | US-10-282-122A-46664 Sequence 46664, A |
| 24         | 245.5 | 7.6         | 724    | 4     | US-10-282-122A-45795 Sequence 45795, A |
| 25         | 242.5 | 7.5         | 793    | 4     | US-10-282-122A-60689 Sequence 23, Appl |
| 26         | 242.5 | 7.5         | 793    | 6     | US-11-045-004-23 Sequence 23, Appl     |
| 27         | 238.5 | 7.4         | 245    | 4     | US-10-333-002-2 Sequence 2, Appl       |

|    |       |     |      |   |                                        |
|----|-------|-----|------|---|----------------------------------------|
| 28 | 232   | 7.2 | 2032 | 3 | US-09-071-035-458 Sequence 458, App    |
| 29 | 232   | 7.2 | 2032 | 3 | US-09-071-035-462 Sequence 462, App    |
| 30 | 232   | 7.2 | 2032 | 3 | US-09-071-035-466 Sequence 466, App    |
| 31 | 232   | 7.2 | 2032 | 4 | US-10-206-576-458 Sequence 458, App    |
| 32 | 232   | 7.2 | 2032 | 4 | US-10-206-576-462 Sequence 462, App    |
| 33 | 232   | 7.2 | 2032 | 4 | US-10-206-576-466 Sequence 466, App    |
| 34 | 232   | 7.2 | 2032 | 5 | US-10-912-362-458 Sequence 458, App    |
| 35 | 232   | 7.2 | 2032 | 5 | US-10-912-362-462 Sequence 462, App    |
| 36 | 232   | 7.2 | 2032 | 5 | US-10-912-362-466 Sequence 466, App    |
| 37 | 226   | 7.0 | 489  | 4 | US-10-282-122A-57763 Sequence 57763, A |
| 38 | 217.5 | 6.7 | 621  | 3 | US-09-071-035-468 Sequence 468, App    |
| 39 | 217.5 | 6.7 | 621  | 4 | US-10-206-576-468 Sequence 468, App    |
| 40 | 217.5 | 6.7 | 621  | 5 | US-10-912-362-468 Sequence 468, App    |
| 41 | 206   | 6.4 | 1103 | 4 | US-10-661-809-9 Sequence 9, Appl       |
| 42 | 202.5 | 6.3 | 402  | 5 | US-10-472-928-780 Sequence 780, App    |
| 43 | 193.5 | 6.0 | 940  | 6 | US-11-045-004-39 Sequence 39, Appl     |
| 44 | 189   | 5.9 | 688  | 3 | US-09-071-035-464 Sequence 464, App    |
| 45 | 189   | 5.9 | 688  | 4 | US-10-206-576-464 Sequence 464, App    |

#### ALIGNMENTS

RESULT 1  
US-09-071-035-218  
Sequence 218, Application US/09071035  
Publication No. US20020045737A1  
GENERAL INFORMATION:  
APPLICANT: GAI H. Choi  
TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides  
NUMBER OF SEQUENCES: 496  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Human Genome Sciences, Inc.  
STREET: 9410 Key West Avenue  
CITY: Rockville  
STATE: Maryland  
COUNTRY: USA  
ZIP: 20850  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage  
COMPUTER: HP Vectra 486/33  
OPERATING SYSTEM: MSDOS version 6.2  
SOFTWARE: ASCII Text  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/071,035  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: A. Anders Brooks  
REGISTRATION NUMBER: 36,373  
REFERENCE/DOCKET NUMBER: PB369P2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (301) 309-8504  
TELEFAX: (301) 309-8512  
INFORMATION FOR SEQ ID NO: 218:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 627 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-071-035-218  
Query Match 100.0%; Score 3225; DB 3; Length 627;  
Best Local Similarity 100.0%; Pred. No. 9.6e-188;  
Matches 627; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MKQLKKWYTVSTLLIPLFTSVIGTTTAAEENGESNAQVIHKKKMTDLPDPLIONSG 60  
DB 1 MKQLKKWYTVSTLLIPLFTSVIGTTTAAEENGESNAQVIHKKKMTDLPDPLIONSG 60

```
QY 61 KEMSEFDKXQGLADYTFSIYVNTNEFEYQRAAGASVDAKQAVQSLTPGKPVAGGTTDAN 120
DB 61 KEMSEFDKXQGLADYTFSIYVNTNEFEYQRAAGASVDAKQAVQSLTPGKPVAGGTTDAN 120
QY 121 GNVTVQLPKKONGKQAVVTIKKEPKEGVAAATNMVAPFVYEMIKQTDGSKYKTEBLAV 180
DB 121 GNVTVQLPKKONGKQAVVTIKKEPKEGVAAATNMVAPFVYEMIKQTDGSKYKTEBLAV 180
QY 181 VHIYKPNVANDGSLHVKVGTAEENGALNGAEFVLSKSGSPGTAVYIQGVADGLYTWTT 240
DB 181 VHIYKPNVANDGSLHVKVGTAEENGALNGAEFVLSKSGSPGTAVYIQGVADGLYTWTT 240
QY 241 DKEQAKRPTTGKSYEIGENDPTEAENGTEGELTVKNLEVGSYILLEEVKAPNNAELIENQTK 300
DB 241 DKEQAKRPTTGKSYEIGENDPTEAENGTEGELTVKNLEVGSYILLEEVKAPNNAELIENQTK 300
QY 301 TPFTTEANNQTPVEKTVKNDTSKVDKTPPSLDGKQVAIGEKKYQISVNIPLGIADKEGD 360
DB 301 TPFTTEANNQTPVEKTVKNDTSKVDKTPPSLDGKQVAIGEKKYQISVNIPLGIADKEGD 360
QY 361 ANKYKFNVLVDKHDALTFDNTSGEYAYALYDGVTVIAPENYQVTEQANGFTVAVNPAY 420
DB 361 ANKYKFNVLVDKHDALTFDNTSGEYAYALYDGVTVIAPENYQVTEQANGFTVAVNPAY 420
QY 421 IPTLTTPGGTLKFEVFMHLENEKADPTKGFKNENAVDNGHTDQTPPTVEVVTGSKRFIKVD 480
DB 421 IPTLTTPGGTLKFEVFMHLENEKADPTKGFKNENAVDNGHTDQTPPTVEVVTGSKRFIKVD 480
QY 481 GNVTVQLPKKONGKQAVVTIKKEPKEGVAAATNMVAPFVYEMIKQTDGSKYKTEBLAV 180
DB 481 GNVTVQLPKKONGKQAVVTIKKEPKEGVAAATNMVAPFVYEMIKQTDGSKYKTEBLAV 180
QY 541 LKGYTYLLEETVAPDDVYLLTNRIEFVNVESYGTENLVSPKVPNKHGTLPSITGKG 600
DB 541 LKGYTYLLEETVAPDDVYLLTNRIEFVNVESYGTENLVSPKVPNKHGTLPSITGKG 600
QY 601 IYVYLGSGAVLLLIAGVFPARRRKENA 627
DB 601 IYVYLGSGAVLLLIAGVFPARRRKENA 627

RESULT 2
US-10-206-576-218
; Sequence 218, Application US/10206576
; Publication No. US20030017495A1
; GENERAL INFORMATION:
; APPLICANT: Choi et al.
; TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides
; NUMBER OF SEQUENCES: 497
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-R
; OPERATING SYSTEM: Windows 98
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/206,576
; FILING DATE: 29-Jul-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 09/071,035
; FILING DATE: 1998-05-04
; APPLICATION NUMBER: US 60/046,655
; FILING DATE: 1997-05-16
; APPLICATION NUMBER: US 60/044,031
; FILING DATE: 1997-05-06
```

```
APPLICATION NUMBER: US 60/066,009
FILING DATE: 1997-11-14
ATTORNEY/AGENT INFORMATION:
NAME: Hyman, Mark J.
REGISTRATION NUMBER: 46,789
REFERENCE/DOCKET NUMBER: PB369P1D1
INFORMATION FOR SEQ ID NO: 218:
SEQUENCE CHARACTERISTICS:
LENGTH: 627 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 218:
US-10-206-576-218

Query Match 100.0%; Score 3225; DB 4; Length 627;
Best Local Similarity 100.0%; Pred. No. 9.6e-18; Indels 0; Gaps 0;
Matches 627; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKQLKRWYVSTLLILPLFTSVLGTTTAFABENGSAQLVIHKKKMTDLPDPLIONSG 60
DB 1 MKQLKRWYVSTLLILPLFTSVLGTTTAFABENGSAQLVIHKKKMTDLPDPLIONSG 60
QY 61 KEMSEFDKXQGLADYTFSIYVNTNEFEYQRAAGASVDAKQAVQSLTPGKPVAGGTTDAN 120
DB 61 KEMSEFDKXQGLADYTFSIYVNTNEFEYQRAAGASVDAKQAVQSLTPGKPVAGGTTDAN 120
QY 121 GNVTVQLPKKONGKQAVVTIKKEPKEGVAAATNMVAPFVYEMIKQTDGSKYKTEBLAV 180
DB 121 GNVTVQLPKKONGKQAVVTIKKEPKEGVAAATNMVAPFVYEMIKQTDGSKYKTEBLAV 180
QY 181 VHIYKPNVANDGSLHVKVGTAEENGALNGAEFVLSKSGSPGTAVYIQGVADGLYTWTT 240
DB 181 VHIYKPNVANDGSLHVKVGTAEENGALNGAEFVLSKSGSPGTAVYIQGVADGLYTWTT 240
QY 241 DKEQAKRPTTGKSYEIGENDPTEAENGTEGELTVKNLEVGSYILLEEVKAPNNAELIENQTK 300
DB 241 DKEQAKRPTTGKSYEIGENDPTEAENGTEGELTVKNLEVGSYILLEEVKAPNNAELIENQTK 300
QY 301 TPFTTEANNQTPVEKTVKNDTSKVDKTPPSLDGKQVAIGEKKYQISVNIPLGIADKEGD 360
DB 301 TPFTTEANNQTPVEKTVKNDTSKVDKTPPSLDGKQVAIGEKKYQISVNIPLGIADKEGD 360
QY 361 ANKYKFNVLVDKHDALTFDNTSGEYAYALYDGVTVIAPENYQVTEQANGFTVAVNPAY 420
DB 361 ANKYKFNVLVDKHDALTFDNTSGEYAYALYDGVTVIAPENYQVTEQANGFTVAVNPAY 420
QY 421 IPTLTTPGGTLKFEVFMHLENEKADPTKGFKNENAVDNGHTDQTPPTVEVVTGSKRFIKVD 480
DB 421 IPTLTTPGGTLKFEVFMHLENEKADPTKGFKNENAVDNGHTDQTPPTVEVVTGSKRFIKVD 480
QY 481 GNVTVQLPKKONGKQAVVTIKKEPKEGVAAATNMVAPFVYEMIKQTDGSKYKTEBLAV 180
DB 481 GNVTVQLPKKONGKQAVVTIKKEPKEGVAAATNMVAPFVYEMIKQTDGSKYKTEBLAV 180
QY 541 LKGYTYLLEETVAPDDVYLLTNRIEFVNVESYGTENLVSPKVPNKHGTLPSITGKG 600
DB 541 LKGYTYLLEETVAPDDVYLLTNRIEFVNVESYGTENLVSPKVPNKHGTLPSITGKG 600
QY 601 IYVYLGSGAVLLLIAGVFPARRRKENA 627
DB 601 IYVYLGSGAVLLLIAGVFPARRRKENA 627

RESULT 3
US-10-661-809-13
; Sequence 13, Application US/10661809
; Publication No. US20040101919A1
; GENERAL INFORMATION:
; APPLICANT: HOOK, Magnus
; TITLE OF INVENTION: BIOINFORMATIC METHOD FOR IDENTIFYING SURFACE-ANCHORED PROTEINS FR
; POSITIVE BACTERIA
```



FILE REFERENCE: P07741US01/BAS  
CURRENT APPLICATION NUMBER: US/10/661,809  
CURRENT FILING DATE: 2003-09-15  
PRIOR APPLICATION NUMBER: 60/410303  
PRIOR FILING DATE: 2002-09-13  
NUMBER OF SEQ ID NOS: 24  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO: 13  
LENGTH: 627  
TYPE: PRT  
ORGANISM: Staphylococcus epidermidis  
US-10-661-809-13

Query Match 100.0%; Score 3225; DB 4; Length 627;  
Best Local Similarity 100.0%; Pred. No. 9.6e-188;  
Matches 627; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKQLKKWYTVSTLLILPLFTSVLGTTTAFABENGSAQLVIHKKQMTDLPDPLIONSG 60  
DB 1 MKQLKKWYTVSTLLILPLFTSVLGTTTAFABENGSAQLVIHKKQMTDLPDPLIONSG 60  
QY 61 KEMSEBDKXQGLADVTFSIYNTNFEYBQRAAGASVDAKQAVQSLTPKQVPAQCTTDAN 120  
DB 61 KEMSEBDKXQGLADVTFSIYNTNFEYBQRAAGASVDAKQAVQSLTPKQVPAQCTTDAN 120  
QY 121 GNVTVQLPKQNGKQAVYTIKEEPKEGVAAATNMVAFVYEMIKQTDGSKYKGTBELAV 180  
DB 121 GNVTVQLPKQNGKQAVYTIKEEPKEGVAAATNMVAFVYEMIKQTDGSKYKGTBELAV 180  
QY 181 VHIYPRNVANDGSLHVKVGTAEENGNGAEGFVLSKSESGPGTVKXIQGVQGLYTWTT 240  
DB 181 VHIYPRNVANDGSLHVKVGTAEENGNGAEGFVLSKSESGPGTVKXIQGVQGLYTWTT 240  
QY 241 DKEQAKRFTGKSYEIGENDFTEAENGTELTIVKNLEVSYLEEVKAPNNAELIENQK 300  
DB 241 DKEQAKRFTGKSYEIGENDFTEAENGTELTIVKNLEVSYLEEVKAPNNAELIENQK 300  
QY 301 TPEFTIANNQTPVEKTVKNDTSKVDKTPSLDGKQVAIEBKIKYQISVNIPLGIADKEGD 360  
DB 301 TPEFTIANNQTPVEKTVKNDTSKVDKTPSLDGKQVAIEBKIKYQISVNIPLGIADKEGD 360  
QY 361 ANKYVFNVLVDGHDALTFDNTSGEYAYALYDGDVIAPEMYQVTEQANGFTVAANPAY 420  
DB 361 ANKYVFNVLVDGHDALTFDNTSGEYAYALYDGDVIAPEMYQVTEQANGFTVAANPAY 420  
QY 421 IPTLPGGTLKRVYFMHLEKADPTKGFQNEANVDGHTDDQTPPTVEVVTGGKRFIKVD 480  
DB 421 IPTLPGGTLKRVYFMHLEKADPTKGFQNEANVDGHTDDQTPPTVEVVTGGKRFIKVD 480  
QY 481 GGVTVQALAGASPVVRDQNSDTANLKIDETTRKATWTKABATTTTADGLVDING 540  
DB 481 GGVTVQALAGASPVVRDQNSDTANLKIDETTRKATWTKABATTTTADGLVDING 540  
QY 541 LKRYTYLEETVAPDDYVLLTRIRIEFVNEQSYGTENLVSPKVPNKHKGTLPTSGKG 600  
DB 541 LKRYTYLEETVAPDDYVLLTRIRIEFVNEQSYGTENLVSPKVPNKHKGTLPTSGKG 600  
QY 601 IYVYLGSGAVLLIAGVYFARRRKENA 627  
DB 601 IYVYLGSGAVLLIAGVYFARRRKENA 627

RESULT 4  
US-10-912-362-218  
Sequence 218, Application US/10913362  
Publication No. US20050043528A1  
GENERAL INFORMATION:  
Applicant: Choi et al.  
TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides  
NUMBER OF SEQUENCES: 497  
CORRESPONDENCE ADDRESSES:  
ADDRESSER: Human Genome Sciences, Inc.  
STREET: 14200 Shady Grove Road

CITY: Rockville  
STATE: Maryland  
COUNTRY: USA  
ZIP: 20850  
COMPUTER READABLE FORM:  
MEDIUM TYPE: CD-R  
COMPUTER: Dell Latitude  
OPERATING SYSTEM: Windows 2000  
SOFTWARE: ASCII Text  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/912,362  
FILING DATE: 06-Aug-2004  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 10/206,576  
FILING DATE: 2002-07-29  
APPLICATION NUMBER: US 09/071,035  
FILING DATE: 1998-05-04  
APPLICATION NUMBER: US 60/046,655  
FILING DATE: 1997-05-16  
APPLICATION NUMBER: US 60/044,031  
FILING DATE: 1997-05-06  
APPLICATION NUMBER: US 60/066,009  
FILING DATE: 1997-11-14  
ATTORNEY/AGENT INFORMATION:  
NAME: Hyman, Mark J.  
REGISTRATION NUMBER: 46,789  
REFERENCE/DOCKET NUMBER: PB369P1D2  
INFORMATION FOR SEQ ID NO: 218:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 627 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 218:  
US-10-912-362-218

Query Match 100.0%; Score 3225; DB 5; Length 627;  
Best Local Similarity 100.0%; Pred. No. 9.6e-188;  
Matches 627; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKQLKKWYTVSTLLILPLFTSVLGTTTAFABENGSAQLVIHKKQMTDLPDPLIONSG 60  
DB 1 MKQLKKWYTVSTLLILPLFTSVLGTTTAFABENGSAQLVIHKKQMTDLPDPLIONSG 60  
QY 61 KEMSEBDKXQGLADVTFSIYNTNFEYBQRAAGASVDAKQAVQSLTPKQVPAQCTTDAN 120  
DB 61 KEMSEBDKXQGLADVTFSIYNTNFEYBQRAAGASVDAKQAVQSLTPKQVPAQCTTDAN 120  
QY 121 GNVTVQLPKQNGKQAVYTIKEEPKEGVAAATNMVAFVYEMIKQTDGSKYKGTBELAV 180  
DB 121 GNVTVQLPKQNGKQAVYTIKEEPKEGVAAATNMVAFVYEMIKQTDGSKYKGTBELAV 180  
QY 181 VHIYPRNVANDGSLHVKVGTAEENGNGAEGFVLSKSESGPGTVKXIQGVQGLYTWTT 240  
DB 181 VHIYPRNVANDGSLHVKVGTAEENGNGAEGFVLSKSESGPGTVKXIQGVQGLYTWTT 240  
QY 241 DKEQAKRFTGKSYEIGENDFTEAENGTELTIVKNLEVSYLEEVKAPNNAELIENQK 300  
DB 241 DKEQAKRFTGKSYEIGENDFTEAENGTELTIVKNLEVSYLEEVKAPNNAELIENQK 300  
QY 301 TPEFTIANNQTPVEKTVKNDTSKVDKTPSLDGKQVAIEBKIKYQISVNIPLGIADKEGD 360  
DB 301 TPEFTIANNQTPVEKTVKNDTSKVDKTPSLDGKQVAIEBKIKYQISVNIPLGIADKEGD 360  
QY 361 ANKYVFNVLVDGHDALTFDNTSGEYAYALYDGDVIAPEMYQVTEQANGFTVAANPAY 420  
DB 361 ANKYVFNVLVDGHDALTFDNTSGEYAYALYDGDVIAPEMYQVTEQANGFTVAANPAY 420  
QY 421 IPTLPGGTLKRVYFMHLEKADPTKGFQNEANVDGHTDDQTPPTVEVVTGGKRFIKVD 480  
DB 421 IPTLPGGTLKRVYFMHLEKADPTKGFQNEANVDGHTDDQTPPTVEVVTGGKRFIKVD 480

Qy 481 GDTATQALAGASFVVRDQNSDTANYLKIDETTKATWTKAEATFTTTADGLVDING 540  
|  
Db 481 GDTATQALAGASFVVRDQNSDTANYLKIDETTKATWTKAEATFTTTADGLVDING 540  
Qy 541 LKGYTYLLEETAPDDYVLLTNRIEFVNEOSYGTENIVSPKVPNKHKGTLPSTGGKG 600  
|  
Db 541 LKGYTYLLEETAPDDYVLLTNRIEFVNEOSYGTENIVSPKVPNKHKGTLPSTGGKG 600  
Qy 601 IYVYLSGAVLLIAGVFAARRKENA 627  
|  
Db 601 IYVYLSGAVLLIAGVFAARRKENA 627  
RESULT 5  
US-09-071-035-220  
; Sequence 220, Application US/09071035  
; Publication No. US20020045737A1  
; GENERAL INFORMATION:  
; APPLICANT: Gai H. Choi  
; TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides  
; NUMBER OF SEQUENCES: 496  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Human Genome Sciences, Inc.  
; STREET: 9410 Key West Avenue  
; CITY: Rockville  
; STATE: Maryland  
; COUNTRY: USA  
; ZIP: 20850  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage  
; COMPUTER: HP Vectra 486/33  
; OPERATING SYSTEM: MSDOS version 6.2  
; SOFTWARE: ASCII Text  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/071,035  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: A. Anders Brookes  
; REGISTRATION NUMBER: 36,373  
; REFERENCE/DOCKET NUMBER: PB369P2  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (301) 309-8504  
; TELEFAX: (301) 309-8512  
; INFORMATION FOR SEQ ID NO: 220:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 560 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-09-071-035-220  
Query Match 89.7%; Score 2892; DB 3; Length 560;  
Best Local Similarity 100.0%; Pred. No. 1,5e-167;  
Matches 560; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 33 BEENGSAQLVIHKKQNTDLPDLIIONSGKSEFPDKYQGLADVTESIVNTEFEYQRA 92  
|  
Db 1 BEENGSAQLVIHKKQNTDLPDLIIONSGKSEFPDKYQGLADVTESIVNTEFEYQRA 60  
Qy 93 GASVAAKQAVOSLTPGKVAAGTTDANGNVVQLPKKONGDAYVTIIEEPKEGVNAT 152  
|  
Db 61 GASVAAKQAVOSLTPGKVAAGTTDANGNVVQLPKKONGDAYVTIIEEPKEGVNAT 120  
Qy 153 NNVAAPVYEMIKQTDGSKYGTIELAVVHIYPKNVVANDGSLHVKVGTAENEGNGAE 212  
|  
Db 121 NNVAAPVYEMIKQTDGSKYGTIELAVVHIYPKNVVANDGSLHVKVGTAENEGNGAE 180  
Qy 213 FVLSKESGSPGTVKYIQGVKGLYTWTTDKQAKRPITGSKSYIEIGENDFTEANGTGELT 272

Db 181 FVLSKESGSPGTVKYIQGVKGLYTWTTDKQAKRPITGSKSYIEIGENDFTEANGTGELT 240  
|  
Qy 273 VGNLEVGSTYILEEVAKPNNALIEHOTKPPFIIEANNOVPVKTKYNDPSKYDKTTPSLD 332  
|  
Db 241 VGNLEVGSTYILEEVAKPNNALIEHOTKPPFIIEANNOVPVKTKYNDPSKYDKTTPSLD 300  
Qy 333 GKDVAIGEKIKYQISVINIPLGIADKEGDANKYKKNLVKDAALTFDQVTSGEYAYALY 392  
|  
Db 301 GKDVAIGEKIKYQISVINIPLGIADKEGDANKYKKNLVKDAALTFDQVTSGEYAYALY 360  
Qy 393 DGDVIAIPENYQVTEQANGFTVAVNPAYIPILTPGGTLKFVYFMHLINEKADPTKGFKNEA 452  
|  
Db 361 DGDVIAIPENYQVTEQANGFTVAVNPAYIPILTPGGTLKFVYFMHLINEKADPTKGFKNEA 420  
Qy 453 NYDNGHTDDQPEPTVEVVTGGRFKVGDVATQALAGASFVVRDQNSDTANYLKIDET 512  
|  
Db 421 NYDNGHTDDQPEPTVEVVTGGRFKVGDVATQALAGASFVVRDQNSDTANYLKIDET 480  
Qy 513 TKAATWTKAEATFTTTADGLVDITGKYGTYLLEETAPDDYVLLTNRIEFVNEOS 572  
|  
Db 481 TKAATWTKAEATFTTTADGLVDITGKYGTYLLEETAPDDYVLLTNRIEFVNEOS 540  
Qy 573 YGTENIVSPKVPNKHKGT 592  
|  
Db 541 YGTENIVSPKVPNKHKGT 560

RESULT 6  
US-10-206-576-220  
; Sequence 220, Application US/10206576  
; Publication No. US20030017495A1  
; GENERAL INFORMATION:  
; APPLICANT: Choi et al.  
; TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides  
; NUMBER OF SEQUENCES: 497  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Human Genome Sciences, Inc.  
; STREET: 9410 Key West Avenue  
; CITY: Rockville  
; STATE: Maryland  
; COUNTRY: USA  
; ZIP: 20850  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: CD-R  
; COMPUTER: Dell latitude  
; OPERATING SYSTEM: Windows 98  
; SOFTWARE: ASCII Text  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/10/206,576  
; FILING DATE: 29-Jul-2002  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 09/071,035  
; FILING DATE: 1998-05-04  
; APPLICATION NUMBER: US 60/046,655  
; FILING DATE: 1997-05-16  
; APPLICATION NUMBER: US 60/044,031  
; FILING DATE: 1997-05-06  
; APPLICATION NUMBER: US 60/066,009  
; FILING DATE: 1997-11-14  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Hyman, Mark J.  
; REGISTRATION NUMBER: 46,789  
; REFERENCE/DOCKET NUMBER: PB369P1D1  
; INFORMATION FOR SEQ ID NO: 220:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 560 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; SEQUENCE DESCRIPTION: SEQ ID NO: 220:

US-10-206-576-220

Query Match 89.7%; Score 2892; DB 4; Length 560;  
Best Local Similarity 100.0%; Pred. No. 1.5e-167;  
Matches 560; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 33 EENGSAQVLIHKKTDLDPDLIIONSGKMSFPDKYQGLADVTFSIYVNTFEYFORA 92
   |||||
DB 1 EENGSAQVLIHKKTDLDPDLIIONSGKMSFPDKYQGLADVTFSIYVNTFEYFORA 60
   |||||
QY 93 GASVDAKAQAVOSLTPGKPAQGTDTANGNVTVLPKQNGKDAVYTIKEEPGEVVAAT 152
   |||||
DB 61 GASVDAKAQAVOSLTPGKPAQGTDTANGNVTVLPKQNGKDAVYTIKEEPGEVVAAT 120
   |||||
QY 153 NMVAPPVYEMIKQTDGSKYGTBEELAVVHIYPKNVANDGSLHVKKVGTAEENGANGAE 212
   |||||
DB 121 NMVAPPVYEMIKQTDGSKYGTBEELAVVHIYPKNVANDGSLHVKKVGTAEENGANGAE 180
   |||||
QY 213 FVYSKSGSPGTIVKIQGVKDGLYTWTDPKQAKRPTGKSYEIGENDFTEANGTGEIT 272
   |||||
DB 181 FVYSKSGSPGTIVKIQGVKDGLYTWTDPKQAKRPTGKSYEIGENDFTEANGTGEIT 240
   |||||
QY 273 VKNLEVGSIYLBKPAKPNNAELIENQTKPTIIEANNQTPVEKTVKNDTSKYDKTTPSLD 332
   |||||
DB 241 VKNLEVGSIYLBKPAKPNNAELIENQTKPTIIEANNQTPVEKTVKNDTSKYDKTTPSLD 300
   |||||
QY 333 GQDVAIGEKIKYQISVNIPLGIADKGDANKYKFNLDKHDALTFPDNVTSGEYVYALY 392
   |||||
DB 301 GQDVAIGEKIKYQISVNIPLGIADKGDANKYKFNLDKHDALTFPDNVTSGEYVYALY 360
   |||||
QY 393 DDDTYIAPENYQVTEQANGFTVAVNPAYIPTLTGGLTKFYVFMHLEKADPTKGRKNA 452
   |||||
DB 361 DDDTYIAPENYQVTEQANGFTVAVNPAYIPTLTGGLTKFYVFMHLEKADPTKGRKNA 420
   |||||
QY 453 NVNDGHTDDQTPPYEVVTVGKRFIKVDGDTATQALAGSFVVRQNSDTANYLIKIDET 512
   |||||
DB 421 NVNDGHTDDQTPPYEVVTVGKRFIKVDGDTATQALAGSFVVRQNSDTANYLIKIDET 480
   |||||
QY 513 TKAATWVKTKAATFTTTADGLVDITGLKGYTYLBEFVAPDDVYLLNRIEFVNEQS 572
   |||||
DB 481 TKAATWVKTKAATFTTTADGLVDITGLKGYTYLBEFVAPDDVYLLNRIEFVNEQS 540
   |||||
QY 573 YGTENLVSPKVPKHKGT 592
   |||||
DB 541 YGTENLVSPKVPKHKGT 560
   |||||
```

RESULT 7  
US-10-912-362-220

Sequence 220, Application US/10912362  
Publication No. US20050043528A1  
GENERAL INFORMATION:  
APPLICANT: ChOI et al.  
TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides  
NUMBER OF SEQUENCES: 497  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Human Genome Sciences, Inc.  
STREET: 14200 Shady Grove Road  
CITY: Rockville  
STATE: Maryland  
COUNTRY: USA  
ZIP: 20850  
COMPUTER READABLE FORM:  
MEDIUM TYPE: CD-R  
COMPUTER: Dell Latitude  
OPERATING SYSTEM: Windows 2000  
SOFTWARE: ASCII Text  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/912,362  
FILING DATE: 06-Aug-2004  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 10/206,576  
FILING DATE: 2002-07-29

APPLICATION NUMBER: US 09/071,035  
FILING DATE: 1998-05-04  
APPLICATION NUMBER: US 60/046,655  
FILING DATE: 1997-05-16  
APPLICATION NUMBER: US 60/044,031  
FILING DATE: 1997-05-06  
APPLICATION NUMBER: US 60/066,009  
FILING DATE: 1997-11-14  
ATTORNEY/AGENT INFORMATION:  
NAME: Hyman, Mark J.  
REGISTRATION NUMBER: 46,789  
REFERENCE/DOCKET NUMBER: PB369PID2  
INFORMATION FOR SEQ ID NO: 220:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 560 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 220:  
US-10-912-362-220

Query Match 89.7%; Score 2892; DB 5; Length 560;  
Best Local Similarity 100.0%; Pred. No. 1.5e-167;  
Matches 560; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 33 EENGSAQVLIHKKTDLDPDLIIONSGKMSFPDKYQGLADVTFSIYVNTFEYFORA 92
   |||||
DB 1 EENGSAQVLIHKKTDLDPDLIIONSGKMSFPDKYQGLADVTFSIYVNTFEYFORA 60
   |||||
QY 93 GASVDAKAQAVOSLTPGKPAQGTDTANGNVTVLPKQNGKDAVYTIKEEPGEVVAAT 152
   |||||
DB 61 GASVDAKAQAVOSLTPGKPAQGTDTANGNVTVLPKQNGKDAVYTIKEEPGEVVAAT 120
   |||||
QY 153 NMVAPPVYEMIKQTDGSKYGTBEELAVVHIYPKNVANDGSLHVKKVGTAEENGANGAE 212
   |||||
DB 121 NMVAPPVYEMIKQTDGSKYGTBEELAVVHIYPKNVANDGSLHVKKVGTAEENGANGAE 180
   |||||
QY 213 FVYSKSGSPGTIVKIQGVKDGLYTWTDPKQAKRPTGKSYEIGENDFTEANGTGEIT 272
   |||||
DB 181 FVYSKSGSPGTIVKIQGVKDGLYTWTDPKQAKRPTGKSYEIGENDFTEANGTGEIT 240
   |||||
QY 273 VKNLEVGSIYLBKPAKPNNAELIENQTKPTIIEANNQTPVEKTVKNDTSKYDKTTPSLD 332
   |||||
DB 241 VKNLEVGSIYLBKPAKPNNAELIENQTKPTIIEANNQTPVEKTVKNDTSKYDKTTPSLD 300
   |||||
QY 333 GQDVAIGEKIKYQISVNIPLGIADKGDANKYKFNLDKHDALTFPDNVTSGEYVYALY 392
   |||||
DB 301 GQDVAIGEKIKYQISVNIPLGIADKGDANKYKFNLDKHDALTFPDNVTSGEYVYALY 360
   |||||
QY 393 DDDTYIAPENYQVTEQANGFTVAVNPAYIPTLTGGLTKFYVFMHLEKADPTKGRKNA 452
   |||||
DB 361 DDDTYIAPENYQVTEQANGFTVAVNPAYIPTLTGGLTKFYVFMHLEKADPTKGRKNA 420
   |||||
QY 453 NVNDGHTDDQTPPYEVVTVGKRFIKVDGDTATQALAGSFVVRQNSDTANYLIKIDET 512
   |||||
DB 421 NVNDGHTDDQTPPYEVVTVGKRFIKVDGDTATQALAGSFVVRQNSDTANYLIKIDET 480
   |||||
QY 513 TKAATWVKTKAATFTTTADGLVDITGLKGYTYLBEFVAPDDVYLLNRIEFVNEQS 572
   |||||
DB 481 TKAATWVKTKAATFTTTADGLVDITGLKGYTYLBEFVAPDDVYLLNRIEFVNEQS 540
   |||||
QY 573 YGTENLVSPKVPKHKGT 592
   |||||
DB 541 YGTENLVSPKVPKHKGT 560
   |||||
```

RESULT 8  
US-10-661-809-19  
Sequence 19, Application US/10661809  
Publication No. US20040101919A1  
GENERAL INFORMATION:  
APPLICANT: HOOK, Magnus

```

; TITLE OF INVENTION: BIOINFORMATIC METHOD FOR IDENTIFYING SURFACE-ANCHORED PROTEINS FR
; FILE REFERENCE: P07741US01/BAS
; CURRENT APPLICATION NUMBER: US/10/661,809
; CURRENT FILING DATE: 2003-09-15
; PRIOR APPLICATION NUMBER: 60/410303
; PRIOR FILING DATE: 2002-09-13
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 19
; LENGTH: 625
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-10-661-809-19

Query Match
Best Local Similarity 75.8%; Score 2446; DB 4; Length 625;
Matches 462; Conservative 64; Mismatches 95; Indels 2; Gaps 1;

1 MKQLKKWYTVSTLLILPLFTSVLGTTTAPFAENGESQQLVIHKKKMTDLPDPLIONSG 60
1 MKNHKINVMGLVLELLPLTNSFGAKKVPFAEET--AAQVILHKKKMTDLPDPLIONSG 58
61 KEMSEFDKYOGLADYTPSIVYNTNEFEQRAAGASVDAKQAVOSLTPGKPVAGCTTDAN 120
59 KEMSEFDYOGLADISFSYVNTQGFYQGRDKGASVDAKQAVOSLTPGTPVASGTTDAD 118
121 GNVYVQLPKQKGNKAVNTIKKEPKEGVVAATMNVVAPVYEMIKQTDGSKYKGBELAV 180
119 GNVVLSLPKQKGNKAVNTIKKEPKQGVSAANMVLAFVYEMIKQADSYKGBELDT 178
181 VAHYEPKNVANDGSLHKKVGTAEENGGLNGAEFVVISKSESGPGTVYIKQVADGLYMTT 240
179 IHLVYKNTVANDGTLKVTIKGTAEENGGLNGAEFVVISKSESGPGTVYIKQVADGLYMTT 238
241 DKEQAKRPTGSKSYEIGENDPTEAENGTELTVKNLEVGSYLIEVKAPNNAELIENQTK 300
239 DQTKAKHFTGHSYDIGNNDFABASIEKGQLIYNHLEVGKYNLEEVKAPDNAEMIEKQTI 298
301 TPFTLEANNQTPVEKTVKNDTSKVDKTPSLDQKVAIGEKIKYQISVNIPLGIADKED 360
299 TPFTLEANSQTPVEKTIKNDTSKVDKTPQLNGKQVAIGEKIKYQISVNIPLGIADKED 358
361 ANKYKFNVLVDGDAALTFDNTSGEYAYALYDGDVIAPENYQVTEQANGFTVANVPAY 420
359 QNKYTFPLKIDHDAALTFDNDSSGTYAYALYDKNKEIDPVNYSVTEQDGFVSDVPNY 418
421 IPTLTGPGTLKVFYFMHLEKADPTKGFKNKANVNDGHTDQTPPTVEVVTGAKRPIKYD 480
419 IPSLTPGGLTKVFYFMHLEKADPTKGFSGNQANVNDGHTNDQTPPSVDVVTGAKRPIKYD 478
481 GDNVTAOLAGASFVVRDQNSDPTANYLKIDETTKAATWTKAEATTTTADGLVDING 540
479 GNVTDQTLTAGAEFVVRDQSDPTAKYLSIDPSTKAVSWSAKESAVFTTTSNGILDVIG 538
541 LKYGTYVEEYVAPDDYVLLTNRIEFVNEQSGYGTENLVSPKYPNKKGLTPTSGKG 600
539 LKYGTYVEEYVAPDDYVLLTNRIEFVNEQSGYGTENLVSPKYPNKKGLTPTSGKG 598
601 IYVYIGAGVALLIAGVTFARRR 623
599 IYVYIGAGVALLIAGVTFARRR 621

RESULT 9
US-10-661-809-11
; Sequence 11, Application US/10661809
; Publication No. US20040101919A1
; GENERAL INFORMATION:
; APPLICANT: HOOK, Magnus
; TITLE OF INVENTION: BIOINFORMATIC METHOD FOR IDENTIFYING SURFACE-ANCHORED PROTEINS FR
; FILE REFERENCE: P07741US01/BAS
```

```

; CURRENT APPLICATION NUMBER: US/10/661,809
; CURRENT FILING DATE: 2003-09-15
; PRIOR APPLICATION NUMBER: 60/410303
; PRIOR FILING DATE: 2002-09-13
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 11
; LENGTH: 476
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-10-661-809-11

Query Match
Best Local Similarity 13.6%; Score 438.5; DB 4; Length 476;
Matches 166; Conservative 70; Mismatches 211; Indels 193; Gaps 22;

8 WYTSVTLILPLFTSVLGTTTAPFAENGESQQLVIHKKKMTDLPDPLIONSGKEMSE 65
6 WLSICVMLLALFGFSQ-----QALAEASQASQVTLHKLPLPDGQLPEQO-QNTGEEGTL 59
66 FDYQGLADYTPSIVYNTNEFEQRAAGASVDAKQ--AVOSLTPGKPVAGCTTDA---- 119
60 LQNYGLNDVITYQVYDVTDPFYQLRSEKTVQEAQRQLAETGATNRKRPFAEDKTQTINGE 119
120 NGNVTVOLPKK---QNGDAVYTIKEBPKEGVAAATMNVVAPVYEMIKQTDGSKYKYTE 176
120 DGVSFSLASKDSQGRDRAVLFVEAEAPVYKASNLVVLFPVQDPQGS----- 170
177 ELAVVHYIPKNVANDGSLHKKVGTAEENGGLNGAEFVVISKSESGPGTVYIKQVADGLY 236
171 -LTHILHYPKN----- 180
227 TWTTDKEQAKRPTGSKSYEIGENDPTEAENGTELTVKNLEVGSYLIEVKAPNNAELIE 296
181 -----BENAYDL----- 187
297 NQTKPTLEANNQTPVEKTVKNDTSKVDKTPSLDQKVAIGEKIKYQISVNIPLGIAD 356
188 -----PFLKTVLDKQGFNQ-----GEHINQVLTQIPANILG 221
357 KEGDANKYKFNVLVDGDAALTFDNTSGEYAYALYDGDVIAPENYQVTEQANGFTVANV 416
222 -----YQEFPLSKADQTLTLTL-----LPESIEVYKAGTVYTG--GYTLTQKHGFTLDF 268
417 NPAYIPLTLPGGTLKVFYFMHLEKADPTKGFKNKAN--VDNGHTDQTPPTVEVVTGAKR 475
269 SIKDLQNA-NQTMVTSQMLKTAEPDTAINNEGQVLTQKGT---LTKRATVARTGKS 324
476 FIKVDGDTATQALAGASFVVRDQNSDPTANYLKIDETTKAATWTKAEATTTTADGL 535
325 FVKVDSE-NAKITLPEAVFIVNQAGEYLN-----ETANGYMQKEKALAKKFTSNQAGE 378
536 VDTGLKTYVEEYVAPDDYVLLTNRIEFVNEQSGYGTENLVSPKYPNKKGLTPTSGKG 592
379 FSVKGLKQGYFLEBISAPKGYLNLQTEIPFTVGNKSAVYNGQRTAPLHVINKKVESGF 438
593 LPSTG-GKGIYVYIGAGVALLIAGV-----YFARRRKE 625
439 LPKTNBRSIMULTING---LLIIGVYIWLFTQKQKGE 474

RESULT 10
US-10-661-809-17
; Sequence 17, Application US/10661809
; Publication No. US20040101919A1
; GENERAL INFORMATION:
; APPLICANT: HOOK, Magnus
; TITLE OF INVENTION: BIOINFORMATIC METHOD FOR IDENTIFYING SURFACE-ANCHORED PROTEINS FR
; FILE REFERENCE: P07741US01/BAS
; CURRENT APPLICATION NUMBER: US/10/661,809
; CURRENT FILING DATE: 2003-09-15
; PRIOR APPLICATION NUMBER: 60/410303
```

PRIOR FILING DATE: 2002-09-13  
NUMBER OF SEQ ID NOS: 24  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 17  
LENGTH: 473  
TYPE: PRT  
ORGANISM: Staphylococcus epidermidis  
US-10-661-809-17

Query Match 12.2%; Score 395; DB 4; Length 473;  
Best Local Similarity 24.4%; Pred. No. 9, 6e-16;  
Matches 156; Conservative 79; Mismatches 200; Indels 204; Gaps 26;

QY 4 LKKV-WYWTST-LILLPLFVSIGTTTAFENSGSADLVHKKMTD--LP--DPLQ 57  
DB 1 MKKLGMLSCFLPLFKPAFTOVA-----TETETEMVOITLHKLLPFGQLPKNHP--- 51  
QY 58 NSGKMSSEPKYQGLADVTFSIYVNTNEFYEORAAAGASVDAKQ--AVQSLTPGKPAQ 115  
DB 52 NQGEKALLQYRGINGVTFQYVDYDSDFYHLREKGYTEBAQAEIAKNGASSGMTAA 111  
QY 116 TT---DANGVTVOL--PKKQNGDAVYTIKEBPKEGV--AATMVVAFPYEMIKQTD 168  
DB 112 TTTTLNNEBGIASFSLAAKQDEKRDYALFIESKYEVVKEKAEMVVLVPHGQ----- 166  
QY 169 GSKYKYTEELAVVHIYPKRVVANDSLHVKYKGTAEENGLAEPVISESGSPGVAKI 228  
DB 167 ---NNOQLSTIHLIYPKN----- 180  
QY 229 QGVKDLVYTWTDKQAKRFTGSKSYEIGENDPTEAENGTEGLTVNLEVGSYILEVVA 288  
DB 181 ---ENDYPD----- 187  
QY 289 PNNAELIENQYTPTEIANNQTPVEKTVKNDTSKVDKTPSLDGDVAIGEKIKQISV 348  
DB 188 ---PPEKYLEBPPN-----DFTIGEKITYSLHT 213  
QY 349 NPLIGIADKESGANKYVKNLVKDDAALTF--DNTSGEYAVALYDGGTVIAPENYQT 406  
DB 214 TLPVILD-----YQFELSDSADALTFPLPNSLTSSNGEKLTEG-----FVH 258  
QY 407 EQANGFTVAVNPAYIPTLT--PGGTLKFYFMHLEKADPTGKPKNEAVNDGHTDDQP 464  
DB 259 KKHPEFDVLF--IPSLKFAVAGKGLTISYQQLSTAAKEINNGLDGF--GVST 313  
QY 465 PTEVAVTGKRFIKYDGDVTAQALAGASFVVRDQNSDTANYIKIDETTKAATWYTKAE 524  
DB 314 KKVSVYTGSKQKQVKIETN--KPKRLAGAVFLIKNK--AGNYIQ--QTANGYKMTKNESD 367  
QY 525 ATFTTTADGLVDIGLAKGYTYLEBTVAPDDYVLLTNRIEFVNVNQSIGTENVLSPK 584  
DB 368 ALHLISDKNGAPSIGLKTGSYRLKRIEAPSGTILSETETPTTIS--TFLSEDKADSLIK 426  
QY 585 VPKKRGKT---LPSTGSKGIYVYLGSAVLLIAGVYFA 620  
DB 427 VVKKENSRPRLPKTN-----ETKNTLLGVGVNFA 457

RESULT 11  
US-09-071-035-274  
Sequence 274, Application US/09071035  
Publication No. US20020045737A1  
GENERAL INFORMATION:  
APPLICANT: Gili H. Choi  
TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides  
NUMBER OF SEQUENCES: 496  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Human Genome Sciences, Inc.  
STREET: 9410 Key West Avenue  
CITY: Rockville  
STATE: Maryland  
COUNTRY: USA  
ZIP: 20850

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage  
COMPUTER: HP Vectra 486/33  
OPERATING SYSTEM: MSDOS version 6.2  
SOFTWARE: ASCII Text  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/071,035  
FILING DATE:

CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:

ATTORNEY/AGENT INFORMATION:  
NAME: A. Anders Brooks  
REGISTRATION NUMBER: 36,373  
REFERENCE/DOCKET NUMBER: PB369P2  
TELEPHONE: (301) 309-8504  
TELEFAX: (301) 309-8512  
INFORMATION FOR SEQ ID NO: 274:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 430 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-071-035-274

Query Match 10.8%; Score 347; DB 3; Length 430;  
Best Local Similarity 24.0%; Pred. No. 7e-13;  
Matches 142; Conservative 64; Mismatches 200; Indels 186; Gaps 19;

QY 8 WYTVSTLILLPLFVSIGTTTAFENSGSADLVHKKMTD--LPDPLQNSGKMS 65  
DB 6 WLSICMLALFQFSQ-----QALBAQASQVQVTLHKLLPFGQLPEQO--QNTGEGTL 59  
QY 66 PDYQGLADVTFSIYVNTNEFYEORAAAGASVDAKQ--AVQSLTPGKPAQGTDA----- 119  
DB 60 LQNTGRLNDVYQVVDVDPFYQLASBEGTVQEAQRLAETGATNRKPIABDKTYQINB 119  
QY 120 NGNVTVOLPKK---QNGKDAVYTIKEBPKEGVAAATMVVAFPYEMIKQTDGSKYKTE 176  
DB 120 DGVSVFSLASKDSQGRDRAVLFVEABAPENVKASNLVILPVDPQGS----- 170  
QY 177 ELAVVHIYPKRVVANDSLHVKYKGTAEENGLAEPVISESGSPGVAKIYQGVSDGLY 236  
DB 171 -LTHIHLIYPKN----- 180  
QY 237 TWTTDKQAKRFTGSKSYEIGENDPTEAENGTEGLTVKNLEVGSYILEBKAPNNAELIE 296  
DB 181 ---ENAYDL----- 187  
QY 297 NQKTPTEIANNQTPVEKTVKNDTSKVDKTPSLDGDVAIGEKIKQISVNIPLGIAD 356  
DB 188 ---PPEKTVLDKQGFNQ-----GEHINQGLTTQIPANILG 221  
QY 357 KEGDANKYKRVLVKDKDAALTFDNTSGEYAVALYDGGTVIAPENYQVTEQANGTVAV 416  
DB 222 ---YQFELSDSKADTTLTL-----LPESIEVAVAGTIVT--GYTLTKKHGHTLDP 268  
QY 417 NPAYIPTLTGPGTLKFYFMHLEKADPTGKPKNEAN--VDNGHTDDQPPTVEVTVGGR 475  
DB 269 SIKDLQNR--NOTMTVSYQMLREKTAEBPTAINNEGQVLTGHT--LTKRATVTRGGS 324  
QY 476 PIKVDGVTATQALAGASFVVRDQNSDTANYIKIDETTKAATWYTKAEATFTTTADGL 535  
DB 325 FVKVDSR--NAKITLPEAVFIYVNGQGEYLN-----ETANGYKQEKALAKKFTSNQGE 378  
QY 536 VDTGLKGYTYLEBTVAPDDYVLLTNRIEFVNVNQSIGTENVLSPKRVN 587  
DB 379 FSVKXKKWPPVLLGRN-----LCTKRLSSESNNSFPYGGKFLCNERTTN 423

```

RESULT 12
US-10-206-576-274
/ Sequence 274, Application US/10206576
/ Publication No. US20030017495A1
/ GENERAL INFORMATION:
/ APPLICANT: Choi et al.
/ TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides
/ NUMBER OF SEQUENCES: 497
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Human Genome Sciences, Inc.
/ STREET: 9410 Key West Avenue
/ CITY: Rockville
/ STATE: Maryland
/ COUNTRY: USA
/ ZIP: 20850
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: CD-R
/ COMPUTER: Dell Latitude
/ OPERATING SYSTEM: Windows 98
/ SOFTWARE: ASCII Text
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/10/206,576
/ FILING DATE: 29-Jul-2002
/ CLASSIFICATION: <Unknown>
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 09/071,035
/ FILING DATE: 1998-05-04
/ APPLICATION NUMBER: US 60/046,655
/ FILING DATE: 1997-05-16
/ APPLICATION NUMBER: US 60/044,031
/ FILING DATE: 1997-05-06
/ APPLICATION NUMBER: US 60/066,009
/ FILING DATE: 1997-11-14
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Hyman, Mark J.
/ REGISTRATION NUMBER: 46,789
/ REFERENCE/DOCKET NUMBER: PB369PDI1
/ INFORMATION FOR SEQ ID NO: 274:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 430 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ SEQUENCE DESCRIPTION: SEQ ID NO: 274:
US-10-206-576-274

Query Match 10.8%; Score 347; DB 4; Length 430;
Best Local Similarity 24.0%; Pred. No. 7e-13; Indels 186; Gaps 19;
Matches 142; Conservative 64; Mismatches 200;

QY 8 WYTVSTLLILPLFTSVLGTTFAPAEENGESAOIVHKKKQTD--LPPDLIONSGKEMSE 65
D  W I S C I M L A L R G F S Q -----Q A L A S Q S V Q V T L H K L F P D Q L P E Q Q -Q N T G E B G T L 59
QY 66 P K I Q G L A V T S I Y V T N E F Y E O R A A G A S V D A K Q -A V O S L T P G K P V A G T T D A --- 119
D  I O N Y R G L I N V T Y Q V V D T P F Q L R S E G K T V O E A Q R O L A E T G A T R R K P I A E D K Q T I N G E 119
QY 120 N G N V T Q L P K K ---Q N G K D A V Y T I K E E P K G G V A A T M V V A F P V E M I K O T D S Y K Y G T E 176
D  120 D G V S S L S K S Q S Q D K A V L F V E A E P E V Y K E K A S H L V I L F P V O D P Q G S ----- 170
QY 177 E L A V V I Y P K N V A N D G S L H F K V K V G T A E N G L N G A E F V I S K E G S P G T V Y I Q G V K G L Y 236
D  171 -L T H I L Y P K N -----E E N A Y D L----- 180
QY 237 T Y T T D K E Q A K R P I T G S Y E I G E N D F T E A N G T G E L T V K N L E V G S Y I L E E V A P N N A E L I E 296
D  181 -----B E N A Y D L----- 187
QY 297 N O T K P T E I A N N Q T E V K T K N D S K V D K T T P S L D G K D V A I G K I K Y Q I S V N I P L G I A D 356

```

```

Db      188 -----PLEKTVLDKXGQFNQ-----GEHINVLTTQIPANILG 221
OY      357 KEGDANKYKFNLDVKHDAALTFDNTVSGEYAYALYDQTVIAPENYQVTEQANGFTAV 416
Db      222 -----YQFRISDRAKDTTLT---LPESIEVKVAGKTVTT--GYTLTKHGHTLDF 268
OY      417 NPAYIFPLTGTGTLTKFYFMHLNEKADPTKGFNMEAN-VDNGHTDQTPPTPEVVTGGR 475
Db      269 SINDLQNEFA-NQMTVSYGKRELEKTAEPDTAINNEGVLTDKHT--LTKRAIVRTGGS 324
OY      476 FIKVDGVATVQALAGASFVVRDQNSDTANYLKIDETTPAATWVKTAETFTTAQGL 535
Db      325 FVKVDSE-NAKITLPEAVFIVKQAGEYLN-----ETANGYWKQEKALAKKFTSNGE 378
OY      536 VDLTGKYGTVYLEETVAPDDVYLLNRIEFVNVNQSXTENTLVSEKAVPN 587
Db      379 FSVAGKRWPLVLLGRN-----LTKRLSSESNNRNSFYGKKFLCNERTTN 423

RESULT 13
US-10-912-362-274
; Sequence 274, Application US/10912362
; Publication No. US20050043528A1
; GENERAL INFORMATION:
; APPLICANT: Choi et al.
; TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides
; NUMBER OF SEQUENCES: 497
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Human Genome Sciences, Inc.
; STREET: 14200 Shady Grove Road
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-R
; COMPUTER: Dell Latitude
; OPERATING SYSTEM: Windows 2000
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/912,362
; FILING DATE: 06-Aug-2004
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 10/206,576
; FILING DATE: 2002-07-29
; APPLICATION NUMBER: US 09/071,035
; FILING DATE: 1998-05-04
; APPLICATION NUMBER: US 60/046,655
; FILING DATE: 1997-05-16
; APPLICATION NUMBER: US 60/044,031
; FILING DATE: 1997-05-06
; APPLICATION NUMBER: US 60/066,009
; FILING DATE: 1997-11-14
; ATTORNEY/AGENT INFORMATION:
; NAME: Hyman, Mark J.
; REGISTRATION NUMBER: 46,789
; REFERENCE/DOCKET NUMBER: PB369PID2
; INFORMATION FOR SEQ ID NO: 274:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 430 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 274:
US-10-912-362-274

Query Match      10 8; Score 347; DB 5; Length 430;
Best Local Similarity 24.0%; Pred. No. 7e-13;
Matches 142; Conservative 64; Mismatches 200; Indels 186; Gaps 19;

8 WYVSTLLILPLFTSVLTGTTTAAFAENESAOVLVHKRMTD--LPDPLIONSGKEMSE 65
::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::

```

Db 6 WLSICMLLALFPGSQ-----QALBAASQASVQVTLHLKLLFPDGLPEQO-QNTGEBGL 59  
Qy 66 PDYQGLADPTESYNNMTEFEYQRAAGSVDAAQO-AVQSLTRPKPAQSTTDA----- 119  
Db 60 LQNYGLADVTYQVYDVDPFYQLRSEGTVEQARQALBETGATNRKPIAEDKTQTNGE 119  
Qy 120 NGNTVAVOLPKK---ONGKDAVYTIKEEPKEGVAAATNMVAFPVYEMIKQTGSGYKYE 176  
Db 120 DGVSFSLASKDSQGRDRAVLFVEBAAPVAVKEKASNLVVIIPVDQPOQS----- 170  
Qy 177 ELAVVHIYPKNVANDGSLHVKKVGTAENEGANGAEFVISKSEBPGTVKYIQGVKDGLY 236  
Db 171 -LTHLHLYPKN----- 180  
Qy 237 TWTTEQAKRITTKSYEIGENDPTBAENGTELTVKYLEVGSYILLEEVKAPNNAELIE 296  
Db 181 -----EENAYDL----- 187  
Qy 297 NQTKPTTEANNQTPVEKTVKNDTSKVDTTPSLDQKQVAILGEKIKYQISVNIPLGIAD 356  
Db 188 -----PPEKTVLDKQGGFNO-----GEHINYQLTQIPANILG 221  
Qy 357 KEGANKYKFNLDVKGDAALTFEDVTSGEYAVLYDGDVTIAPENYQVTEQANGFTAV 416  
Db 222 -----VQEFRISDKADTTLTL-----LPESIEVKVAGKTVTT--GYTLTQKHGFTLDF 268  
Qy 417 NPAYIFLTPGSLKFVYFMHLNEKADPTKGRKNEAN-VDNGHTDDQTPPYEVYTGGR 475  
Db 269 SKDQONRA-NQTMVVSQMRLEKTAEPRTAINNBOQLTDTGTT--LTKRAVTRGGS 324  
Qy 476 FIKVGDVATQALAGASPVVRDQNSDPTANYLKIDETTKATWTKAEATFTTTADGL 535  
Db 325 FVKVUSE-NAKTLTBEAVIVKQAGEYLN-----ETANGYRQKEKALAKKFTSNQGE 378  
Qy 536 VDTGLKGTGYLLEETVAPDDYVLLTNRIEFVNVESYGTENTLVSPEKVPN 587  
Db 379 FSVKXKKMPVLLGRN-----LCTKRLSSSNRNSFYGKKFLCNERTTN 423

RESULT 14  
US-10-333-002-16  
; Sequence 16, Application US/10333002  
; Publication No. US20040071729A1  
; GENERAL INFORMATION:  
; APPLICANT: Adderson, Elisabeth  
; APPLICANT: Bohnsack, John  
; TITLE OF INVENTION: Group B Streptococcus Polypeptides Nucleic Acids and Therapeutic  
; FILE REFERENCE: 2511-1-001 (SJ-0039)  
; CURRENT APPLICATION NUMBER: US/10/333,002  
; PRIOR FILING DATE: 2003-07-08  
; PRIOR APPLICATION NUMBER: PCT/US01/24795  
; PRIOR FILING DATE: 2001-08-08  
; PRIOR APPLICATION NUMBER: US 09/634,341  
; NUMBER OF SEQ ID NOS: 37  
; SOFTWARE: Patent in version 3.1  
; SEQ ID NO 16  
; LENGTH: 502  
; TYPE: PRT  
; ORGANISM: Streptococcus agalactiae  
US-10-333-002-16

Query Match 10.4%; Score 334; DB 4; Length 502;  
Best Local Similarity 26.3%; Pred. No. 5,3e-12;  
Matches 150; Conservative 79; Mismatches 223; Indels 118; Gaps 26;

Qy 100 KOAVQSL-----TRGPVAGSTTDANGNVVQPKQONGDAVYTIKEEPKEGV 148  
Db 3 KKMISLVLASLAFKAUSPVTPPIAFA--ETGTTVQDTQKATYKAKVDAEIDNAN 60  
Qy 149 VAATNMVAFPVYEMIKQTGSGYKYEELAVVHIYPKNVANDGSLHVKKVGTAE-NEG 207

Db 61 VDSNMDGA--SYLIPQKEAEYKASTDPSNLF-----TTTTNGRTVYTKDITASANE- 112  
Qy 208 LINGAFVLSKSGSGYTKYIQGVNDGLYTTWTDKEQARFTGKSYELGENDPTEAEN- 266  
Db 113 -----IATVA-----KSIISANTPV-STVYESND 136  
Qy 267 GTBELTVKULEVGSYILLEEVKAPNNAELIENQTKPTTEANNQTPVEKTVKNDTSKVDK 326  
Db 137 GTEVINVS--QGYTYVS--STVNGGAVIMTSVTP-----NATIHKE--NTDAIWDG 184  
Qy 327 TTPSLDQKQVAILGEKIKYQISVNIPLGIADKEGKANKYKFNLDVKGDAALTFEDVTSGE 386  
Db 185 GSKTVQDKTVSVGDVTKKTLITYKANVNHGTE----KVQYVIKDTPBASVVD-LNEGS 239  
Qy 387 YAVLYDGDVTI-----APENYQVTEQANGFTAVANPVIPTLTGCG----- 428  
Db 240 YEVTTTDSGNITTLTQSEKATYKKNLLENNPTTIT--DWAATNTPTGNTONGAND 297  
Qy 429 -----TLKFVYFMHLNEKADPTKG-----FNGEANVDNGHTDDQTPPYEVYTGGRPI 477  
Db 298 FPKYGINITTYTTLGVLKSGAKPGSADLPENNTINATINPNSNDPQGVTVRQGITIK 357  
Qy 478 KVDGVTATQALAGASPVVRDQNSDPTANYLKIDETTKATWTKAEATFTTTADGLVD 537  
Db 358 KIDGSTKA--SLQGAIFVLKNA---TGQFLNFDTN--VEM-GTBAANTEYTTGADGIT 410  
Qy 538 ITGLKGTGYLLEETVAPDDYVLLTNRIEFVNVESYGT--ENLVSPEKVPKHKGT--LP 594  
Db 411 ITGLKEGYLYLVEKAPLGVNLLDSQKVIADGATDTTNSDNLNLPVEN-NKGTELP 469  
Qy 595 STGKGIVYLAGSAGVLLIAGVFPARRK 624  
Db 470 STGIGITTFITIIIGALVIGAGIVVARRR 499

RESULT 15  
US-09-071-035-276  
; Sequence 276, Application US/09071035  
; Publication No. US20020045737A1  
; GENERAL INFORMATION:  
; APPLICANT: Gili H. Choi  
; TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides  
; NUMBER OF SEQUENCES: 496  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Human Genome Sciences, Inc.  
; STREET: 9410 Key West Avenue  
; CITY: Rockville  
; STATE: Maryland  
; COUNTRY: USA  
; ZIP: 20850  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage  
; COMPUTER: HP Vectra 486/33  
; OPERATING SYSTEM: MSDOS version 6.2  
; SOFTWARE: ASCII Text  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/071,035  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: A. Anders Brookes  
; REGISTRATION NUMBER: 36,373  
; REFERENCE/DOCKET NUMBER: PB369P2  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (301) 309-8504  
; TELEFAX: (301) 309-8512  
; INFORMATION FOR SEQ ID NO: 276:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 384 amino acids  
; TYPE: amino acid









Db 312 KNIGDKDB---NNGETIATAKHDITANLITTYFTDY--VDRENSVOMGINSIYMDADT 366  
Qy 96 VDAKAQVOSLTPGKVAAGTTPDANGNVYQLP-----KKONG 133  
Db 367 IPVSKNDVFP-----NVTIGNTTKTTANIQYDPYVNEKNSIGAFETESHVGNKEP 421  
Qy 134 KDAVYTIKEEPKEGVAAATNMV-----AFP-----YEMIK--QTDGSYK 172  
Db 422 GYTKQITTYNPBENSILNAKLKVAQVHSSYPNNIGQINDVDIKIYQVPKGTTLKNKGD 481  
Qy 173 YGTEELA-VVHIYPRKVVANDGSLHVKVGTAE--GLNGAEFVLSKSGSPGVYKIQ 229  
Db 482 VNTKELTDVITNOGLQKITGDNNSAVIDFGNDSAVVWMTFQYTNSE--SPTLVQMA- 539  
Qy 230 GYKQGLYTTTDEKQAKRPIYKSYEIGENDFTEANGELTVKLEFGSYLLEEVKAP 289  
Db 540 -----TLSTGNKSVSTGNALGFTNNQ-----SGAGQOEYK--IGNVYMEDTNK 583  
Qy 290 NNAEL-----IENQTKPTIEANNQTPV-EKTYKNDTSKVDKTPSLDGKDAIGEKIKY 344  
Db 584 GVOELDEKGVGNTTVF--DNNNTKVGBAVTKEGSLIPLPNGD-----Y 630  
Qy 345 QISV-NIPLG---IADKEGDANKYVKNLVKHDALTPDNVTS--GEYAYALYDGTVI 398  
Db 631 RVEFNSLPKGYEVTPEKQGN-NBELDSNGLSVITVNGKONSADLGIVKPKYNLGDYVW 689  
Qy 399 ABEYV-----QYTEQANGTVAV--NPAYIPLLTGGGLTKFYFMHLEKADPTG-EX 449  
Db 690 EDTNKGIDODDEKIGSGVTATLKONGVNLKTVITDADGKYF-----TDLNGNMYK 742  
Qy 450 NEAVNNGHTDDQTPTEVTVTGKRFIVDGDVTAQLAGA-----492  
Db 743 VEFTPEEGT-----PT--TVTSGDIEKDSNGLITTVINGADNNTLDSGFKTPKYNL 795  
Qy 493 -SPVVRDQNSDYANLIKIDETTKAATWVTKAE-----ATFTTADGLVDITGLKYG 545  
Db 796 GNYVVEDTKDG--KQDSTEKIGSGVTATLKNGEVLQTTKTDKQGYQPTGLENG 851  
Qy 546 YLLE-ETVAPDDVYLTNRIFPVANEQSYGTENLV-----PEKVPNKHGKGLPS- 595  
Db 852 YVVEEET--PSGV-----TPQVGSGETDEGIDSGTSTTGVIKDKNDITISGFYK 900  
Qy 596 -TGSGKIYVY 604  
Db 901 PTTYNLGDYVW 910

RESULT 4  
US-10-471-571A-2278  
Sequence 2278, Application US/10471571A  
Publication No. US20060115490A1  
GENERAL INFORMATION:  
APPLICANT: CHIRON SPA  
TITLE OF INVENTION: STAPHYLOCOCCUS AUREUS PROTEINS AND NUCLEIC ACIDS  
FILE REFERENCE: P026927MO  
CURRENT APPLICATION NUMBER: US/10/471,571A  
CURRENT FILING DATE: 2003-09-12  
PRIOR APPLICATION NUMBER: GB-0107661.1  
PRIOR FILING DATE: 2001-03-27  
NUMBER OF SEQ ID NOS: 5642  
SOFTWARE: SeqMin99, version 1.03  
SEQ ID NO 2278  
LENGTH: 2478  
TYPE: PRT  
ORGANISM: Staphylococcus aureus  
FEATURE:  
NAME/KEY: MISC FEATURE  
LOCATION: (1) - (2478)  
OTHER INFORMATION: FmbB protein [Staphylococcus  
US-10-471-571A-2278

Query Match 4.6%; Score 149.5; DB 6; Length 2478;  
Best Local Similarity 21.7%; Pred. No. 0.32;

Matches 136; Conservative 79; Mismatches 237; Indels 175; Gaps 34;  
Qy 58 NSGKSEFQDKQGLADVTSIYNTNEFYEQRAAAS-VDAKAQVOSL--TPGKPVQ 114  
Db 1676 NAVKTFSEYK--DALAKIEDANAKVNEADNSASTSSIEAKQGLAEIKQIADQVNO 1734  
Qy 115 GTT-----DANGNVYQLPKONGDAVYTIKEEPKEGVAAATNMVAVPVE 162  
Db 1735 ATSKQDIEVQIHNDLDNIDYITIPGKESATTDLYAVADQKNNISLQTN-ATQEKQ 1793  
Qy 163 MIKQTDGSKYKTEEL-----AVVH-----IYKVVANDGSLHYK 198  
Db 1794 AIKQVQNVQTLAESINGGVNDGDVDDALQGGAAIDAQVATVFK--ANQ-AIEYK 1849  
Qy 199 KYGTAEENGLNABFVLSKSGSPGVYK-----QGVQGLYTTTDEKQAKRPIYKGS 253  
Db 1850 ABDYKESIDQSOLTAEBKTE-ALMIKQITQAKQGITDA--TTTAEVEKAK---AQG 1902  
Qy 254 YEIGEN---DTEAENGTEL--TVKNLEVGSYI-----LEEVKAPNNA--ELIENQTK 301  
Db 1903 LEAFDNIQIDSTBKQKATIELETDQIAGVNVANADATTEKEKFTNALBEDLSKATE- 1961  
Qy 302 PTTIEANNQTPVEK--TVKN--DTSKVDKTPSLDGKDV-AIGEKIKYQISVNIPLGIA 355  
Db 1962 ----DISQTTNABEATVNGSALBQKQRIINPEYKKNLBALIREVVKQIEI--IKNA 2014  
Qy 356 DEKGDANKYVKNL-----VDKHDALTPDNVTSGEVAYALYDGTVIABEYVYEQ 408  
Db 2015 DADSAKEIARFDLGRYFDRFADKLDQYT-----NAEVAEL 2051  
Qy 409 ANGFTVAVPAVYIPLLTPGGLTKFYFMHLEKADPTGFKNE--ANYDNGHTDDQ- 463  
Db 2052 QN-VITPAEALVPQNDP-----DANDTNGIINDNDATANSNAAATPEPTQ 2097  
Qy 464 PPTVEVYTGKRFK-----VDGVTATQALAGSPVVRDQNSDYANLIKIDETTK 514  
Db 2098 PNVSEFTANGKADASPTTENNDAATGETTATSATDADNDKQANNNS-----VDASTN 2152  
Qy 515 AATW-----VTKAEATTTTADGLVDITGLKGYIIEYVAPDDVYLTNRIFPVN 569  
Db 2153 SPTMNDVTSKEVSESTNNGT-----DKPVTETDN--ATPAESTTN 2192  
Qy 570 EOSYGTENLVSEPEKVPNKHGKGLPST 596  
Db 2193 NNSTTTATV---ENAPGSTATAFTT 2215

RESULT 5  
US-10-471-571A-860  
Sequence 860, Application US/10471571A  
Publication No. US20060115490A1  
GENERAL INFORMATION:  
APPLICANT: CHIRON SPA  
TITLE OF INVENTION: STAPHYLOCOCCUS AUREUS PROTEINS AND NUCLEIC ACIDS  
FILE REFERENCE: P026927MO  
CURRENT APPLICATION NUMBER: US/10/471,571A  
CURRENT FILING DATE: 2003-09-12  
PRIOR APPLICATION NUMBER: GB-0107661.1  
PRIOR FILING DATE: 2001-03-27  
NUMBER OF SEQ ID NOS: 5642  
SOFTWARE: SeqMin99, version 1.03  
SEQ ID NO 860  
LENGTH: 2186  
TYPE: PRT  
ORGANISM: Staphylococcus aureus  
FEATURE:  
NAME/KEY: MISC FEATURE  
LOCATION: (1) - (2186)  
OTHER INFORMATION: Mtp protein [Staphylococcus  
US-10-471-571A-860

Query Match 4.5%; Score 146; DB 6; Length 2186;  
Best Local Similarity 20.3%; Pred. No. 0.44;

Matches 140; Conservative 89; Mismatches 269; Indels 192; Gaps 31;

|    |      |                                                                  |      |
|----|------|------------------------------------------------------------------|------|
| Qy | 23   | VLIGTTTFAEENGSSAOLVHKKQWTDLPDLJONSGKEMSEFPXQGLADVFSINY           | 82   |
| Dd | 1559 | SIKOTPHATVDBELDEANOLISDTLKOAOOE---IENTODAAVTD-----VBNQ           | 1644 |
| Qy | 83   | TNEFYEO-----RAAGASV-----DAKQAVOSLTPEKFAOQT-----                  | 116  |
| Dd | 1645 | TIKAEIQIKPKVRRKRAALDSTIEENNKQOLDAIRTLTPTQOERDVALIDLINKIVNTIKN    | 1704 |
| Qy | 117  | -----TDANGNVTOQLPKQKNGKADAVYTIKEEPKSGVAAITMVAFAVY                | 161  |
| Dd | 1705 | DIQNKTNABVDRIETTDGNDNKVILPKVQO-----VKPAAQSV-----GVAAEQN          | 1751 |
| Qy | 162  | EMIKOTDSSYKXGTEELAVVHIYPKPV---VANDSLHVKKVCGAEBEGNCAEFVSKSEG      | 220  |
| Dd | 1752 | ALIDPSDLSTE---BERLIAKHVEGALNQAIDQINHAKTKQVQVNDSTIN---QONISK---   | 1805 |
| Qy | 221  | SPGIYKVIQGVYDGLYTTTDDKEQAKRPITGKSYEI---GENDFTAEENGTEGLTVKNLEVG   | 279  |
| Dd | 1806 | -----IKPATTYATKALQOQIONATYKINKINKANNEATDEE-----QNIATA            | 1847 |
| Qy | 280  | SYILEEYKAPNNALIEENQTKPFTI---EANNQOTVEKTVKNDTSKVDKTPSPSLDGKQVA    | 337  |
| Dd | 1848 | QVEKELIQAQOIASAVTNADVAYILHBEKKEIREIBEYINRKASAREQTLTLFPDCKQA      | 1907 |
| Qy | 338  | IGEKIRKYOIS-----VNIPLGIADKE---GDANKYVENLVKDDAALTFPDNV            | 382  |
| Dd | 1908 | IEANIQATVEERNSILAOLONIYDTALQIIDDRESNAGVDKTASLNTQTIDHLDV----      | 1962 |
| Qy | 383  | TSGEVAVALYOGDYIAPENVOYVEQANGFIYAVANVPATIPULTPBGTLKFVFMHINKKA     | 442  |
| Dd | 1963 | -----HPIKKPDKEKTIINDLAKVATLVONRYKRSN-----RNKA                    | 1997 |
| Qy | 443  | DPYTGFKPN---EANVNDGHTDDQTPTEVEVVTGSKREIKYDGDVTAOTALAGASFVVRDONS  | 501  |
| Dd | 1998 | DALKAIATLKLQMDBELKTARITNADVDVAVL---KRNVALSDIEAV-----LREKENS      | 2047 |
| Qy | 502  | DTANVLKIDDETTKAAITWYKTAKEATTTFTTTADGLVDITGL---KY---GTYILEEYVAADD | 556  |
| Dd | 2048 | -----LIRIDNITAQ---QTYAKFAFAIA-----TEQOLAKVVLIDQYAADNRMIMDEBATLND | 2097 |
| Qy | 557  | YVLLTNRIEFVUNE-----OSYGTENVLVPEPK-----PNKHKGT-----LPSY           | 596  |
| Dd | 2098 | ---IKQHQFIYDDELAIKLPRAEANKVSPKEIQAPAKVCTPIKKEETHESKVEKELEPNT     | 2154 |
| Qy | 597  | GKGGIYVYLSGAVALLIAGYFPAARRKEN                                    | 626  |
| Dd | 2155 | GSEGMDDLPLKEFA---LITGAALLLARRITTN                                | 2182 |

## RESULT 6

```

US-10-471-571A-4496
? Sequence 4496, Application US/10471571A
? Publication No. US20060115490A1
? GENERAL INFORMATION:
? APPLICANT: CHIRON SPA
? TITLE OF INVENTION: STAPHYLOCOCCUS AUREUS PROTEINS AND NUCLEIC ACIDS
? FILE REFERENCE: P026927W0
? CURRENT APPLICATION NUMBER: US/10/471,571A
? CURRENT FILING DATE: 2003-09-12
? PRIOR APPLICATION NUMBER: GB-0107661.1
? PRIOR FILING DATE: 2001-03-27
? NUMBER OF SEQ ID NOS: 5642
? SOFTWARE: Seqwin99, version 1.03
? SEQ ID NO 4496
? LENGTH: 9535
? TYPE: PRT
? ORGANISM: Staphylococcus aureus
? FEATURE:
? NAME/KEY: MISC FEATURE
? LOCATION: (1)..(9535)
? OTHER INFORMATION: hypothetical protein

```

US-10-471-571A-4496

|                       |                  |                 |             |              |
|-----------------------|------------------|-----------------|-------------|--------------|
| Query Match           | 4.5%;            | Score 145.5;    | DB 6;       | length 9535; |
| Best Local Similarity | 20.3%;           | Pred. No. 4.2;  |             |              |
| Matches 140;          | Conservative 83; | Mismatches 295; | Indels 173; | Gaps 30      |

|    |      |                                                                |      |
|----|------|----------------------------------------------------------------|------|
| QY | 24   | VLGTTTAAFAENGESAOVLHKKKMDLDPRLQNSGKEMSEBDK---QGLADTYFSIY       | 80   |
| Db | 1436 | VKSNSTINADAVYKQOLYLPYGPROYEHLNQNNGNTTDAINITYFVPSDLVNPTISVG     | 1495 |
| QY | 81   | NVTNE---FYEORAAAGASVDAAKQAVOSLT-PGKPVAGQTTDAN-GNVTQVLEPKQKQKXA | 136  |
| Db | 1496 | NYTNHQVPSSEITNTITANDNGVOSVYVPTNSOQTGVVDNNHQHVSATAPNYSATMK      | 1555 |
| QY | 137  | VYTIKEEPKEGVAAITMVVAFPVYEMIKOTDSYKGTBELAVHVIYPRQVANDGSLH       | 196  |
| Db | 1556 | TINLATDITSGNATTSFNVT-----VKPLRDKRVGTSSTAMPVRJAN-ISMNATVS       | 1608 |
| QY | 197  | VKKVGTAEENGNGAEFVSKS-----EGSPGVXYIOGVKDGJLT                    | 237  |
| Db | 1609 | QADOTTIIN-SLTFTEIYVPRKRSIARASANEITSKYTSVNSYRGNANAVTYVYQDGTIS   | 1667 |
| QY | 238  | WTT--DKEQAKRPITGKSVEIGENDPTEANGTGBELTKNLEVGSIYIEV-----KAP      | 289  |
| Db | 1668 | TVTVPVGVHVPEIYVAISHSHYVQODE-PAGNGSSASDYPEKLSNGSDIADATITWVSGAP  | 1726 |
| QY | 290  | N--NABIEHQTKPTPIEANNQTPUEKTYVNDTSVVDTPTESLQKDVAIIEKIKYOS       | 347  |
| Db | 1727 | NKQDTRIGEDITVTAHILIDGETTPTTKA--TYVAVATP-----KAVFEYA            | 1772 |
| QY | 348  | VNIPL-GIADKEGDANKYKFNLVDKHDAALTFDNTSGEYAVAYDGDVTIAPENYQVT      | 406  |
| Db | 1773 | RGVLYPQVSD-----MYDAQOYKVPVNNSMS                                | 1798 |
| QY | 407  | EQANGPFAVNAVPIYPLP-----GTLAKFYVFMHLNE-----KADPTKGFQK           | 450  |
| Db | 1799 | TNAQ---HNHFQVFGYGENKDVGISTRLIKVTDNQBEDLTILSKYKPPDPRIIDAN       | 1854 |
| QY | 451  | EANVDNGHTD-----DQTPVLEVVTGGRKPIKV-----DGDV                     | 483  |
| Db | 1855 | SVTYKQGLTNQEIKANNVLNNSVKLFKADNRPPLANTNTIHSQGSFVSVTVSADALPNGGI  | 1914 |
| QY | 484  | TAIYQALA--GASVAVDQNSDTANYK--IDETTLQATWYKTKAEATFTTTAGLVNI       | 538  |
| Db | 1915 | KAKSSISMNNVYTTTQDEHQVVTVTRNBSVDSNDSATVTVPOLOATT-----EBAVPI     | 1969 |
| QY | 539  | ---TGAKYGYTYLEE-----TYA-----PDDY---VLTNRLEFVNEOSYGTETL         | 579  |
| Db | 1970 | KGDGDPFG--HYERFIQNPBHGAITYAMHDSPTTKNVTGNHHTKAVLTVLPNGQG--TRNV  | 2026 |
| QY | 580  | VSPKVPNKRKGTLPSTGKGIVYIGSNAV                                   | 610  |
| Db | 2027 | EVVPKYVPVANAAPASRDVKGQNLITNGTAM                                | 2057 |

## RESULT 7

```

US-10-530-879--4
; Sequence 4, Application US/10530879
; Publication No. US20060140980A1
;
; GENERAL INFORMATION:
; APPLICANT: GUGS, Bengt et al.
; TITLE OF INVENTION: IMMUNIZATION OF NON-HUMAN MAMMALS AGAINST STREPTOCOCCUS EQUI
; FILE REFERENCE: 0825-0173PUS2
; CURRENT APPLICATION NUMBER: US/10/530,879
; CURRENT FILING DATE: 2005-04-11
; PRIOR APPLICATION NUMBER: PCT/SE2003/001587
; PRIOR FILING DATE: 2003-10-10
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
;
; LENGTH: 657
;
; TYPE: PRT
; ORGANISM: Streptococcus equi
;

```



;; TITLE OF INVENTION: LYMPHATIC ENDOTHELIAL GENES  
;; FILE REFERENCE: 28967/39178  
;; CURRENT APPLICATION NUMBER: US/10/505,928  
;; CURRENT FILING DATE: 2004-08-27  
;; PRIOR APPLICATION NUMBER: US 60/363,019  
;; PRIOR FILING DATE: 2002-03-07  
;; NUMBER OF SEQ ID NOS: 866  
;; SOFTWARE: PatentIn 3.2  
;; SEQ ID NO: 449  
;; LENGTH: 3396  
;; TYPE: PRT  
;; ORGANISM: Homo sapiens  
US-10-505-928-449

Query Match 4.2%; Score 135; DB 6; Length 3396;

Best Local Similarity 18.6%; Pred. No. 4.2; Indels 188; Gaps 27;

Matches 113; Conservative 82; Mismatches 223; Indels 188; Gaps 27;

QY 38 SAOLYHHKKMTDLDDPLIONSCKMSEFDKYQGLADVTFSIYNVTNEFEQRAAGASVD 97  
DB 2486 SYMLPLHBEQNKSSPD-----TSTLSNTYS--YKSTDSGSPD 2522  
QY 98 AAKQAVQSLTPGKPYAQGTTDANGVTVOLPKKQNGKDAVYTIKE---EPKEGVAAVN 153  
DB 2523 RFRFEDSTL--KPRRKPKTE---NIIDLDKED--KDLILTITESTIIEILPELTSDKN 2575  
QY 154 MVVAE-----PYVEMI--KOTDSYKKGTEBELAVNHIYKRVAVANGSLHVKV--GTARE 206  
DB 2576 TIIDIDHTKPYEDILGQOTDIDTEPSEP-----HSDNSNDSTOVQSIYENAVNL 2629  
QY 207 GINGAEFVYSKESGPGTV--KYIQGVKGLYTTTDDKEQAKRFITGKSYEIGENDFTEAE 265  
DB 2630 SLTEETF-----EGSADVLASTQATHDSMTY--EDRSQLD----- 2664  
QY 266 NGTGLTVKNLEVGSIIEEVAPNNAE---LIENQTKPTPTIEANOTPEVKTAVKNDT 321  
DB 2665 -----HMGHFTTGIAPSTETELDVILPATISLPIPKSATVIPEIEGIKAEA 2713  
QY 332 SKVD--KTPSLDGDVAIGKI-----KYQ-----ISVNIPLGIADKE 358  
DB 2714 KALDMFESSTLSDQALADQSEIITPLGQFERTQEBYEDKKGAGSFPSPSSGAEBAL 2773  
QY 359 GDANKYKFNLYDKHDAALT-----FDNVTSGEYAVLYDGDVTYIAPENYOV 405  
DB 2774 VDHTEYLSIATTHLMDQSTEVPDWEGSNPRYITDTLAVSTPAKLSQTSPLTIYS 2833  
QY 406 TEQANGFTVAANPAYIP-----TLPQGTILKFVY-----FMELNEKADPT 445  
DB 2834 GSEAGSGHEIPQPSALPGIDGVSSVWSPDSFKFEIHVNIEATPKPSSEBYLHITPPSL 2893  
QY 446 KGFKAENVNDGH-----TDDQTPPTVEVVTGKRFI-----KYDGDVTATQALAGASPV 496  
DB 2894 PDTKLEPSDDQKPELLEEMASPTIELINVEGTEIIODFQNKTDQVSG----- 2942  
QY 497 RDQNSDTANYLKIDETTKAATWTKAEATFTTTADGLVDITGLKY-----GTYYL 548  
DB 2943 -----EAIKMFPIIKT--PEAGTVITTADE--IELEGATQWPHSTASATYGV 2986  
QY 549 EETVAP 554  
DB 2987 EAGVVP 2992

RESULT 10

US-11-256-173-28  
;; Sequence 28, Application US/11256173  
;; Publication No. US20060140979A1  
;; GENERAL INFORMATION:  
;; APPLICANT: University of Sheffield  
;; TITLE OF INVENTION: Antigenic Peptides  
;; FILE REFERENCE: Coxin  
;; CURRENT APPLICATION NUMBER: US/11/256,173  
;; CURRENT FILING DATE: 2005-10-24

;; PRIOR APPLICATION NUMBER: US/10/311,879  
;; PRIOR FILING DATE: 2003-03-18  
;; NUMBER OF SEQ ID NOS: 32  
;; SOFTWARE: PatentIn Ver. 2.1  
;; SEQ ID NO: 28  
;; LENGTH: 2659  
;; TYPE: PRT  
;; ORGANISM: Staphylococcus aureus  
US-11-256-173-28

Query Match 4.2%; Score 134.5; DB 7; Length 2659;

Best Local Similarity 19.1%; Pred. No. 3.2; Indels 155; Gaps 24;

Matches 113; Conservative 73; Mismatches 250; Indels 155; Gaps 24;

QY 81 NVTNEFEQRAAGASVDAAKQAV---QSITPGKPYAQGT-----TDANGVTVOLPKK 130  
DB 1551 NPTMDVNTVNOKAAVSKTKDLDDQQLQRAKTAHTAHTASDLNDQAQKALTOQVNS 1650  
QY 131 QNGKDAVYTIKEPKEGVAAATNMVVAEPVYEMIKQTD-----GSYKYGTE 176  
DB 1651 AQNVQAVNDIKQTTQSLMTAATGLKRGVANHQVQSDNYVNAVDTNKKNDYNNAYNHAND 1710  
QY 177 ELAVVHIIP-----KNVANDGSL-----HVKKVGTARENG 207  
DB 1711 IINGNAQHPVITPPSDVNNALSVTTSKEHALNGEAKLNAKQDANTALGHLNINNAQRON 1770  
QY 208 L-----NGAEFVYSKESGPGTVY---IQGVKGLYTTTDDKEQAKRFITGKSYEIGEN 259  
DB 1771 LOSQINGAHQI-----DAVNTIKQNAATNLNSMGHLROAVADKQDYK-----TE 1815  
QY 260 DPTAEANGTGLTVKNLEVGSIIEEVAPNNALIENQTKPTPTIEANOTPEVKTAVKN 319  
DB 1816 DYAAD-----TAKQNAVNS-----AVSAEFTINQTTNP--TMSYVD-----VNR 1854  
QY 320 DTSKVDKTPSLDGDVAIGEKIKYQISVN--IP-----LGIAKKEGANKYKFNLYD 371  
DB 1855 AHSATYSKNKALNGYEKLAQSTDAARAIIDALPHLNNQKADVKIKINAASIAGVNTVK 1914  
QY 372 KHDAALTFDNVTSGEYAVLYDGDVTYIAPENYQ--VTEQANGFTVAANPAYIPTLPQGT 429  
DB 1915 QQGTPL---NTAMGHLQGAINDBOITLNSQNYQDAPTSKTKTYTNAVQAA----- 1961  
QY 430 LKFVYFMELNEKADPTKGFKNAEVANDGHDTDDQTPPTVEVVTGKRFIKVDG-----D 482  
DB 1962 -----KQILNKSNGQ--KTKQVTEAMNQVSAKN--NLDGTRLLDQAK 2002  
QY 483 VTAATQALAGASFTVADQNSDTANYLKIDETTKAATWTKAEATFTTTADGLVDITGLK 542  
DB 2003 QTAQOOLNNMTHLTTAQKTNLNTQINSCTTAVAGVQTVQSNV-----NTLDDAMNT--LR 2054  
QY 543 YGTYYLAEFTVAPDDYVLLTNRIEFVNVNEQSYGTENLVSPEKVPNKHKGTL 593  
DB 2055 QSIANKDKATKASEDIYDANNNDQKTAYN--NAVAAAEITIIINANSNPEMNPSTI 2104

RESULT 11

US-10-449-902-50884  
;; Sequence 50884, Application US/10449902  
;; Publication No. US20060123505A1  
;; GENERAL INFORMATION:  
;; APPLICANT: National Institute of Agrobiological Sciences.  
;; APPLICANT: Bio-oriented Technology Research Advancement Institution.  
;; APPLICANT: The Institute of Physical and Chemical Research.  
;; APPLICANT: Foundation for Advancement of International Science.  
;; TITLE OF INVENTION: FULL-LENGTH PLANT CDNA AND USES THEREOF  
;; FILE REFERENCE: MOA-A020571-US  
;; CURRENT APPLICATION NUMBER: US/10/449,902  
;; CURRENT FILING DATE: 2003-05-29  
;; PRIOR APPLICATION NUMBER: JP 2002-203269  
;; PRIOR FILING DATE: 2002-05-30  
;; PRIOR APPLICATION NUMBER: JP 2002-383870  
;; PRIOR FILING DATE: 2002-12-11  
;; NUMBER OF SEQ ID NOS: 56791

SOFTWARE: PatentIn Ver. 2.1  
 SEQ ID NO 50884  
 LENGTH: 799  
 TYPE: PRT  
 ORGANISM: Oryza sativa  
 US-10-449-902-50884

Query Match  
 Best Local Similarity 19.9%; Pred. No. 1;  
 Matches 121; Conservative 92; Mismatches 233; Indels 162; Gaps 27;

4.0%; Score 130; DB 6; Length 799;  
 108 PKPVAQGTIDANGN-----VTQLPKQNGKDAVYTIKEPKKGVAAATNVV 156  
 215 PDDCVSR--TTVCNGNSKDKTCSDDYGLNHNKGNLGLPATLAGEKKRNSPPNNGQL 273  
 157 APPVYEMIKQTDGSKYKTEBELAVNHYRK--NVVANDGSLHYKGVTAENEG---LNGA 211  
 274 SASVSGELPPLGPKVHNLNDQLASCDDKPOASVOSANGNLSKQV--TAAANGTVGTLSAK 332  
 212 EKVISKSGSPGVKXIQGVKQGLVTTTDDKEQAKFPIGKSYEIGENDFTBAENGTEL 271  
 333 QTVNVVSSQSSS-----GRRTVLTRQTLASDTRSKATG---QVG-----NASSDQKL 379  
 272 TVKNLEVGSYILEEVKAPNNABELIENQTKPTPTTEAN-----NQTPEKTVKNDTSKYD 325  
 380 TSANNEHSRRI--KISRSDNVKLVSGRPREPQGMLANHLTGALDKTHVDTDEKNASDIN 437  
 336 KTT-----PSLDGKQVAILGEKIKYQISVNIPLG 353  
 438 EKVTCGIQWQLKESTPAHRSTVLQSLRDPMNSNLPFLD-----VKSQISV----- 483  
 354 IADKGDANKYKVFNLV---DKHDAALTFEDNTSGEYALVYDGVIAPEVYQVTEQAN 410  
 484 VEDKSDSOSASRTQIQPSNHNKTAVCSSDTANASD--ACGIANOVLPFGKGRQTSQGE 542  
 543 DSHLVKRDSSGSDQLSSQHPGNVPSRLLTSLSSIDIRAKENKGIKRNV----- 592  
 411 GFTV-----AVNPAYIPTLPFGTLK---FVYFMHLEKADPTKGFKNENYDNGHTDQ 462  
 543 DSHLVKRDSSGSDQLSSQHPGNVPSRLLTSLSSIDIRAKENKGIKRNV----- 592  
 463 TPPTVEVVTGKRFKVDGVATQAL--AGASFVVRDONS--DTANVLIKIDETTKATW 518  
 593 CFPGEBELRPDSKITSVSSPTSSIMCSGPDVLQDSCSAKDQDFI-----SW 643  
 519 VTKAEATFTTTADGLVDITGLKYGTYYLEBTVAPDY-----VLTNRLEFVNBBS 572  
 644 VSECLDEGGETTQSNKSIPT-----LSSTDAWRMYQPPPVFLVHQLTFLVSPYP 695  
 573 YGTTEILVSPKVPN-----KHK-----GTLPSYTGKGIYVYLGSQA 609  
 696 ROLSGHTVG--RIENTMNCCHSPSVSGIANHKPEVWGSCHSVMSTG---YDVPSQSA 750  
 610 VLLIAGV 617  
 751 TLGMIAGM 758

RESULT 12  
 US-11-236-173-15  
 Sequence 15, Application US/11256173  
 Publication No. US20060140979A1  
 GENERAL INFORMATION:  
 APPLICANT: University of Sheffield  
 TITLE OF INVENTION: Antigenic Peptides  
 FILE REFERENCE: toxin  
 CURRENT APPLICATION NUMBER: US/11/256,173  
 CURRENT FILING DATE: 2005-10-24  
 PRIOR APPLICATION NUMBER: US/10/311,879  
 PRIOR FILING DATE: 2003-03-18  
 NUMBER OF SEQ ID NOS: 32  
 SOFTWARE: PatentIn Ver. 2.1  
 SEQ ID NO 15  
 LENGTH: 960  
 TYPE: PRT

ORGANISM: Staphylococcus aureus  
 US-11-256-173-15

Query Match  
 Best Local Similarity 21.4%; Pred. No. 2,8;  
 Matches 146; Conservative 69; Mismatches 265; Indels 202; Gaps 35;

8 WTVSTLLILLPLFSV-----LGTTFAEENGSAQVVIHKKQTTDLPDLIQSGE 62  
 121 WGTOSTITPTTSKPTTBSKSTGKLTVA--NNGVAQL--KPTNSGLYTTVYDKTGKA 175  
 63 MSEFDKYQGLADVTESI-----YVNTFEYQRAAGASVDAKQAVQ-- 104  
 176 TNEVQK-----TAVSKTATLGNQKPYLDVQDYSNGKFGVWKGSDVYVNTAKSVVNN 228  
 105 ---SLTPG-----KPVAGTTDANGVTVQLPKQNGKDAVYTI--KEEPKGV 149  
 229 QESYIKPQTKLVTPWGTSKQVA--GSVSGSGQTFKASQOOIDKSIYLXGSGVNGSGWV 287  
 150 AATNNV-----VAPVYEMIKQTDGSKYKTEBELAVNHYPRN-----VANDGSLHYK 199  
 288 SNAVLDIAKPTPTPKPSTPTNNKLTSSLSNGVQAQINAKONGLFTTVYDKTGKPYKE 347  
 200 V-----GTAEENGAGAFVYSKSGSPGVKTYI--Q-----VQDQ-- 234  
 348 VQKTRAVYTKBASLGNKFLVVDYNSPTLIGVKKQGDVYVNNAKSPVNMQYTVKPGTK 407  
 235 LYT--WTTDKQA-----KRPITGKSYEIGEND--FTEAENGTELTVKNLEVGSYI 282  
 408 LYSVPMGTYKQBAAGVSGTGNQTFKATKQOOIDKSIYLEGTVNGKSGWVSKAVLAVPAAP 467  
 283 LEEVKAAPNNABELIENQTKPT-----IANN-----QTVVEKTVKNDTSKYDTS 330  
 468 KAAVQKPTAAVAVYVTKPTQTVSKIAQVAPNNTGIPASVYEKTAAGAKYADPT--F 525  
 331 LDGKDAIGEK--IKYQISVNIPLGIADKEDANKYKVFNLVDKDAALTFEDNTSGEY 387  
 526 YTKERAHENETVYLVNNTSHNIPLG-----WPNVXD-----LANNQL----- 563  
 388 AVALYDGVTVAPENVQTEQANGFTVAVNPAYIPTLPFGTLKVFYFMHLEKADPTG 447  
 564 -----GKEVKTKQKTYVKSNNGL-----SMVPMGT--KQVLIITGNNAIQGTEN 606  
 448 FKNENAVD-----NGHT-----DQOTPT--VEVVGGRFVKVDGVATQALA--G 491  
 607 ATKQVSGKDVLYGTINNRIGVMNAKDLTAPFAVPTISAQ-----DYNTVYIKNG 660  
 492 ASFVVRDONSDFANY--LKI-----DETTKATVYTK-----AEA 525  
 661 NGYVYVTRNSDTAKSLAFNEQPPAVVKEQVYINGQTVYGLKSNGLAMIKSTDLAKEL 720  
 526 TTFPTTTADGLVDITGLKYGTYY 547  
 721 IKYNOTGMALNQVAQIQAGLQY 742

RESULT 13  
 US-11-105-233-158  
 Sequence 158, Application US/11105233  
 Publication No. US20060134653A1  
 GENERAL INFORMATION:  
 APPLICANT: Thialagalingam et al  
 TITLE OF INVENTION: Differential Expression of Genes in MSI  
 FILE REFERENCE: Tumors  
 CURRENT APPLICATION NUMBER: US/11/105,233  
 CURRENT FILING DATE: 2005-04-13  
 NUMBER OF SEQ ID NOS: 202  
 SOFTWARE: FastSeq for Windows Version 4.0  
 SEQ ID NO 158  
 LENGTH: 2355  
 TYPE: PRT  
 ORGANISM: Homo sapiens

US-11-105-223-158

```
Query Match      3.9%; Score 124.5; DB 7; Length 2355;
Beet Local Similarity 19.9%; Pred. No. 11;
Matches 130; Conservative 82; Mismatches 245; Indels 197; Gaps 30;

QY 29 TAFV-EEENGESKQVLIHKKKQTDLPDPLIIONSGKEMSEFDKYQGLADYFISYVNTNEFY 87
DB 879 TIYANEEQNESTPRVVI-QQETGTGRSDVPSRDL-----QFVEVTDVAVTITMTRPP----- 930
QY 88 EORAAASVDAK-----QAVOSLTPG-----KPV- 112
DB 931 ESAVYGVYVDVLPVNLPGHGGRLPISRTFAELVTGLSGVYVYKVFVAVSHGRESKPLT 990
QY 113 AAGTTDANGNTVQVLPKQNGKDAVYTIKEEKEGVATAITNVVAFPVYEMIKOD-----G 169
DB 991 AAGTTLKLAFTNLQF---VNETDSTVLVAKMTPPRAQITGYRLTVGLTRRGQPRQNVGDS 1047
QY 170 SYKYTEELAVVHIIPKRVNVANDGSLHVKVGTAEENGINGAEFYISKEGSPGTVKYIQ 229
DB 1048 VSKYELRLNLPASEETVSLVAIKNGQESFKATGVFTTLQPGSSIPRYNTEVETTLI----- 1103
QY 230 GVKDGLVYTTWTDKEQAKR--FITGKSYEIGENDFTEAENGTELTVKMLEVG-SYTL--- 283
DB 1104 ----VITWT-----PAPRIGFYLGYRPSQSGEAPREVTSDSGSIYVSGILTPGEVEYVYTIQ 1154
QY 284 ----EEVKAPNNALIEHQTKTPT-----IEANNQTPVKEKYKNDTSKYDKTTPSLD 332
DB 1155 VLRDQGEADAP-----IVNKKVTPPLSPPTNLHLEANPDGV-LTVSWES-----TTPDIT 1204
QY 333 GKDVAIIGEIKIKYQISVNIPLGIADKEGDANKYKFNLDKHAALTFDNTVSG-EYAAVL 391
DB 1205 G-----YRITTTPTNG--QQGNSLEEV---VHADQSSCTFQRLSLGLEYNVSV 1247
QY 392 YDGDVIAIPENYQVTEQANGFTVAVNPAYITPLTREGTLKFTYFMHLEKADPTGFKNE 451
DB 1248 Y-----TVKDKESVPIISDITIIIPAVPPPTDLRF-----TN 1277
QY 452 ANVDNGHTDDQTPPVEVNVVVGSKRPKV-----DGDVTAOTALAGASFYVR 497
DB 1278 IGPDMRTWAPPSIDLTNPLVRSFVNKEEDVALSISPSDNAVLTNLPLPTEYVVS 1337
QY 498 ----DONSDT-----ANYLKIDETTKATWKTVAEATP----- 528
DB 1338 VSSVYQEHSTPLRGQKTGLDPSGIDPSDITANSFTVHMLAPATTIGYRIHHPHF 1397
QY 529 -----TTTADGLVDITGLKKGTYYLEETVAPDDYVLLTNRIE--FVNEQS 572
DB 1398 SGRPREDRVPHSRNSITLTNLTPGTEYVVSIVA-----LNGRESPLLIQOQS 1445
```

RESULT 14

```
US-10-471-571A-5118
; Sequence 5118; Application US/10471571A
; Publication No. US20060115490A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SPA
; TITLE OF INVENTION: STAPHYLOCOCCUS AUREUS PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE: P026927WO
; CURRENT APPLICATION NUMBER: US/10/471,571A
; CURRENT FILING DATE: 2003-09-12
; PRIOR APPLICATION NUMBER: GB-0107661.1
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 5642
; SOFTWARE: SeqMan99, version 1.03
; SEQ ID NO 5118
; LENGTH: 1263
; TYPE: PRP
; ORGANISM: Staphylococcus aureus
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (1)..(1263)
; OTHER INFORMATION: autoLysin [Staphylococcus
```

US-10-471-571A-5118

```
Query Match      3.8%; Score 124; DB 6; Length 1263;
Beet Local Similarity 21.4%; Pred. No. 4.8;
Matches 136; Conservative 64; Mismatches 208; Indels 228; Gaps 35;

QY 79 IYVNTNEFEORAAASVDAKQAVOSLTPGKPVAAQGTTDANGVNTVQLPKKONGKDAVY 138
DB 406 LYDLINEXKLLKMGKVAAPWGQSTTPTTPSPK----- 438
QY 139 TKEEPEKQ--VVAATNVVAFPVYEMIKQDGSYKYGTBELAVVHIIPKRVNVAND---G 193
DB 439 TTPSPKPSYTKLTVANNVVA-----QIKPTN-SGLYTT-----VYDKGKATNEVQK 484
QY 194 SLHVKVGTAEENGEL-----NGAEF-----VISSEGPSGTVKYIQGVQDS--LYT 237
DB 485 TPVASKITVLLGQKRYLVQDYNQKFKGVKEGVDVYVNTAASPVVNVSYSIKPQTKLYT 544
QY 238 --WTTDKQAKRFITGKSYEIGENDFTEAENGTELTVK-----NLEVGSYILLEVKAP- 289
DB 545 VPWGTSKQVA-----GVSQSGNQTFKASKQOQIDKSIYLYGSMNGKS 587
QY 290 ---NNALIEHQTKTPTTIEANNQTPVKEKYKNDTSKYDKTTPSLDGDVAIGEIKIKYQI 346
DB 588 GWVSKAYLVDDAKPTP-----TPTPKP-STPTTNKLTIVSSLSNG--VA-----QI 629
QY 347 SV---NIPLGIADKEGDANKYV-----KENLVDKHAALTFDNTVSGEYVA 388
DB 630 NAKKNGLTYYDYDKGKRTKEVQKTFVNTKASLQKGFYLVKQVNSFTLIGWYKQGVY 689
QY 389 YALYDGDVIAIPENYQVTEQANGFTVAVNPAYITPLTREGTLKFTYFMHLEKADPTGKF 448
DB 690 Y-----NNAKSPVNVQTYTVKKGTKLYSV-----PWGTY 719
QY 449 KNEANVDNGHTDDQTPP-----TYEVNG--GKRIRKYDQ---DVTATQ 487
DB 720 KOEAGAVSG-IGNQTFKATKQOQIDKSIYLYGTVNGKSGWSKAYLVAPAKKVAAPQK 778
QY 488 ALAGASPVVRDONSPTVANYLKIDETTKATWKTVAEATTTFTTADGLVDITGLKRYG--T 545
DB 779 TAVKAYITTKQTTQTVS--KIAQYKPNNTGR--ASVYEKTKAK-----NGATYADRT 827
QY 546 YTL--EETVAPDDYVLLTNRIEFV-----VNEQSYG---TTENLVSPDK----- 584
DB 828 FYVTERAHAGNETYVLLNNTSHNIPLGWFNVADLVNQVIGKEVKTQKYTVKSNNGLSM 887
QY 585 VP-----NKKHGTLPSTG-----GKGIYVY 604
DB 888 VFWGTKNQVILLGNNAQGTENATKQVSVGKDVLY 923
```

RESULT 15

```
US-10-471-571A-3542
; Sequence 3542; Application US/10471571A
; Publication No. US20060115490A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SPA
; TITLE OF INVENTION: STAPHYLOCOCCUS AUREUS PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE: P026927WO
; CURRENT APPLICATION NUMBER: US/10/471,571A
; CURRENT FILING DATE: 2003-09-12
; PRIOR APPLICATION NUMBER: GB-0107661.1
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 5642
; SOFTWARE: SeqMan99, version 1.03
; SEQ ID NO 3542
; LENGTH: 1629
; TYPE: PRP
; ORGANISM: Staphylococcus aureus
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (1)..(1629)
; OTHER INFORMATION: hypothetical protein
```





**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.9  
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: July 6, 2006, 02:36:51 ; Search time 50 Seconds  
(without alignments)  
1097.634 Million cell updates/sec

Title: US-10-661-809a-13

Perfect score: 3225

Sequence: 1 MKQLKKWMTVSTLLILPL.....GAVLLIAGVPRARRKENA 627

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 650591 seqs, 87530628 residues

Total number of hits satisfying chosen parameters: 650591

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Issued Patents AA:  
1: /EMC Celerra\_SIDS3/ptodata/2/1aa/5\_COMB.pep:\*  
2: /EMC Celerra\_SIDS3/ptodata/2/1aa/6\_COMB.pep:\*  
3: /EMC Celerra\_SIDS3/ptodata/2/1aa/7\_COMB.pep:\*  
4: /EMC Celerra\_SIDS3/ptodata/2/1aa/H\_COMB.pep:\*  
5: /EMC Celerra\_SIDS3/ptodata/2/1aa/PCUTS\_COMB.pep:\*  
6: /EMC Celerra\_SIDS3/ptodata/2/1aa/RB\_COMB.pep:\*  
7: /EMC Celerra\_SIDS3/ptodata/2/1aa/backfile1.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description         |
|------------|-------|-------------|--------|----|---------------------|
| 1          | 3225  | 100.0       | 627    | 2  | US-09-071-035-218   |
| 2          | 3225  | 100.0       | 627    | 2  | US-10-206-576-218   |
| 3          | 3225  | 100.0       | 659    | 2  | US-09-134-000C-6124 |
| 4          | 2882  | 89.7        | 560    | 2  | US-09-071-035-220   |
| 5          | 2882  | 89.7        | 560    | 2  | US-10-206-576-220   |
| 6          | 2446  | 75.8        | 664    | 2  | US-09-107-532A-7252 |
| 7          | 438.5 | 13.6        | 477    | 2  | US-09-134-000C-6123 |
| 8          | 414.5 | 12.9        | 660    | 2  | US-09-107-532A-6715 |
| 9          | 395   | 12.2        | 508    | 2  | US-09-107-532A-5331 |
| 10         | 347   | 10.8        | 430    | 2  | US-09-071-035-274   |
| 11         | 347   | 10.8        | 430    | 2  | US-10-206-576-274   |
| 12         | 311   | 9.6         | 384    | 2  | US-09-071-035-276   |
| 13         | 311   | 9.6         | 384    | 2  | US-10-206-576-276   |
| 14         | 296.5 | 9.2         | 665    | 2  | US-09-769-787-127   |
| 15         | 290.5 | 9.0         | 341    | 2  | US-09-107-532A-5384 |
| 16         | 232   | 7.2         | 2032   | 2  | US-09-071-035-458   |
| 17         | 232   | 7.2         | 2032   | 2  | US-09-071-035-462   |
| 18         | 232   | 7.2         | 2032   | 2  | US-09-071-035-466   |
| 19         | 232   | 7.2         | 2032   | 2  | US-10-206-576-458   |
| 20         | 232   | 7.2         | 2032   | 2  | US-10-206-576-462   |
| 21         | 232   | 7.2         | 2032   | 2  | US-10-206-576-466   |
| 22         | 232   | 7.2         | 2054   | 2  | US-09-134-000C-6612 |
| 23         | 219.5 | 6.8         | 413    | 2  | US-09-107-532A-6433 |
| 24         | 217.5 | 6.7         | 621    | 2  | US-09-071-035-468   |
| 25         | 217.5 | 6.7         | 621    | 2  | US-10-206-576-468   |
| 26         | 201   | 6.2         | 1154   | 2  | US-09-134-000C-6122 |

|    |       |     |      |   |                     |                   |
|----|-------|-----|------|---|---------------------|-------------------|
| 27 | 197.5 | 6.1 | 1161 | 2 | US-09-327-536-2     | Sequence 2, App11 |
| 28 | 196.5 | 6.1 | 1095 | 2 | US-09-107-532A-3855 | Sequence 3855, Ap |
| 29 | 193.5 | 6.0 | 905  | 2 | US-09-134-000C-4420 | Sequence 4420, Ap |
| 30 | 193.5 | 6.0 | 1112 | 1 | US-08-714-402-2     | Sequence 2, App11 |
| 31 | 190.5 | 5.9 | 910  | 2 | US-09-134-000C-4288 | Sequence 4288, Ap |
| 32 | 190   | 5.9 | 1027 | 2 | US-09-107-532A-6675 | Sequence 6675, Ap |
| 33 | 190   | 5.9 | 1221 | 2 | US-09-107-532A-3959 | Sequence 3959, Ap |
| 34 | 189   | 5.9 | 688  | 2 | US-09-071-035-464   | Sequence 464, App |
| 35 | 189   | 5.9 | 688  | 2 | US-10-206-576-464   | Sequence 464, App |
| 36 | 189   | 5.9 | 1074 | 2 | US-09-071-035-358   | Sequence 358, App |
| 37 | 189   | 5.9 | 1074 | 2 | US-09-071-035-394   | Sequence 394, App |
| 38 | 189   | 5.9 | 1074 | 2 | US-10-206-576-358   | Sequence 358, App |
| 39 | 189   | 5.9 | 1074 | 2 | US-10-206-576-394   | Sequence 394, App |
| 40 | 189   | 5.9 | 1096 | 2 | US-09-134-000C-5764 | Sequence 5764, Ap |
| 41 | 184   | 5.7 | 952  | 2 | US-09-107-532A-4706 | Sequence 4706, Ap |
| 42 | 183.5 | 5.7 | 390  | 2 | US-09-769-787-128   | Sequence 128, App |
| 43 | 183   | 5.7 | 1638 | 2 | US-09-071-035-258   | Sequence 258, App |
| 44 | 183   | 5.7 | 1638 | 2 | US-09-071-035-262   | Sequence 262, App |
| 45 | 183   | 5.7 | 1638 | 2 | US-09-071-035-266   | Sequence 266, App |

## ALIGNMENTS

RESULT 1  
US-09-071-035-218  
Sequence 218, Application US/09071035  
Patent No. 6448043  
GENERAL INFORMATION:  
APPLICANT: G11 H. Choi  
TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides  
NUMBER OF SEQUENCES: 496  
CORRESPONDENCE ADDRESSES:  
ADDRESSER: Human Genome Sciences, Inc.  
STREET: 9410 Key West Avenue  
CITY: Rockville  
STATE: Maryland  
COUNTRY: USA  
ZIP: 20850  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.4MB storage  
COMPUTER: HP Vectra 486/33  
OPERATING SYSTEM: MSDOS version 6.2  
SOFTWARE: ASCII Text  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/071,035  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: A. Anders Brooks  
REGISTRATION NUMBER: 36,373  
REFERENCE/DOCKET NUMBER: PB369P2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (301) 309-8504  
TELEFAX: (301) 309-8512  
INFORMATION FOR SEQ ID NO: 218:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 627 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-071-035-218  
Query Match 100.0%; Score 3225; DB 2; Length 627;  
Best Local Similarity 100.0%; Pred. No. 1.5e+26; Indels 0; Gaps 0;  
Matches 627; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
1 MKQLKKWMTVSTLLILPLFTSVLGTTAFENGESAOVYHKKMTDLPDPLIONSG 60

```

Db      1 MKQLKKWYTVSTLLILPLFTSVLGTTTAPAEENGESAQLVIHKKMTDLPDPLIONSG 60
QY      61 KEMSEFDKYGGLADYTFISIYNTNFEYEQRAAGASVDAKQAVOSLTTPKPAQGTDDN 120
        61 KEMSEFDKYGGLADYTFISIYNTNFEYEQRAAGASVDAKQAVOSLTTPKPAQGTDDN 120
Db      121 GNVTVQLPKQNGKDAVYTIKEEPKEGVAAATNMVAPFVYEMIKQTDSYKGTBEELAV 180
QY      121 GNVTVQLPKQNGKDAVYTIKEEPKEGVAAATNMVAPFVYEMIKQTDSYKGTBEELAV 180
        121 GNVTVQLPKQNGKDAVYTIKEEPKEGVAAATNMVAPFVYEMIKQTDSYKGTBEELAV 180
Db      181 VHIYKPNVANDGSLHVKKVGTAEENGGLNGAEFVLSKESGPGTVKXIQGVNDGLYTWTT 240
QY      181 VHIYKPNVANDGSLHVKKVGTAEENGGLNGAEFVLSKESGPGTVKXIQGVNDGLYTWTT 240
        181 VHIYKPNVANDGSLHVKKVGTAEENGGLNGAEFVLSKESGPGTVKXIQGVNDGLYTWTT 240
QY      241 DKEQAKRFITGKSYEIGENDFTEAENGTELTVKULEVGSYLLEEVKAPNNALLIENQTK 300
        241 DKEQAKRFITGKSYEIGENDFTEAENGTELTVKULEVGSYLLEEVKAPNNALLIENQTK 300
Db      301 TPFTIEANNQTPVEKTVKNDTSKVDKTPPSLDGKQVAIGEKKYQISVNIPLGIADKED 360
QY      301 TPFTIEANNQTPVEKTVKNDTSKVDKTPPSLDGKQVAIGEKKYQISVNIPLGIADKED 360
        301 TPFTIEANNQTPVEKTVKNDTSKVDKTPPSLDGKQVAIGEKKYQISVNIPLGIADKED 360
Db      361 ANKYKFNLVDRHDAALTFDNTVSGEYAYALYDGTVIAPENYQVTEQANGFTVAVNPAY 420
QY      361 ANKYKFNLVDRHDAALTFDNTVSGEYAYALYDGTVIAPENYQVTEQANGFTVAVNPAY 420
        361 ANKYKFNLVDRHDAALTFDNTVSGEYAYALYDGTVIAPENYQVTEQANGFTVAVNPAY 420
Db      421 IPTLPGGTLKRVYFMHINEKADPTKGFKNBANVDNGHTDDQTPPTVEVYVVGKGFIRKD 480
QY      421 IPTLPGGTLKRVYFMHINEKADPTKGFKNBANVDNGHTDDQTPPTVEVYVVGKGFIRKD 480
        421 IPTLPGGTLKRVYFMHINEKADPTKGFKNBANVDNGHTDDQTPPTVEVYVVGKGFIRKD 480
Db      481 GGVTAOTALAGASFVVRDQNSDTANYLKI DETTKAATVTKAABATFTTTADGLVDING 540
QY      481 GGVTAOTALAGASFVVRDQNSDTANYLKI DETTKAATVTKAABATFTTTADGLVDING 540
        481 GGVTAOTALAGASFVVRDQNSDTANYLKI DETTKAATVTKAABATFTTTADGLVDING 540
Db      541 LKGYTYLEETVAPDDYVLLTNRIEFVNEOSYGTTEMLVSPKYPNKHGTLPTSGGKG 600
QY      541 LKGYTYLEETVAPDDYVLLTNRIEFVNEOSYGTTEMLVSPKYPNKHGTLPTSGGKG 600
        541 LKGYTYLEETVAPDDYVLLTNRIEFVNEOSYGTTEMLVSPKYPNKHGTLPTSGGKG 600
Db      601 IYVYLSGAVLLLIAGVYFARRRKENA 627
QY      601 IYVYLSGAVLLLIAGVYFARRRKENA 627
        601 IYVYLSGAVLLLIAGVYFARRRKENA 627
Db      601 IYVYLSGAVLLLIAGVYFARRRKENA 627

RESULT 2
US-10-206-576-218
; Sequence 218 Application US/10206576
; Patent No. 6913907
; GENERAL INFORMATION:
; APPLICANT: Choi et al.
; TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides
; NUMBER OF SEQUENCES: 497
; CORRESPONDENCE ADDRESS:
; ADDRESSSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-R
; OPERATING SYSTEM: Windows 98
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/206,576
; FILING DATE: 29-Jul-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 09/071,035
; FILING DATE: 1998-05-04
; APPLICATION NUMBER: US 60/046,655
; FILING DATE: 1997-05-16
; APPLICATION NUMBER: US 60/044,031
```

```

; FILING DATE: 1997-05-06
; APPLICATION NUMBER: US 60/066,009
; FILING DATE: 1997-11-14
; ATTORNEY/AGENT INFORMATION:
; NAME: Hyman, Mark J.
; REGISTRATION NUMBER: 46,789
; REFERENCE/DOCKET NUMBER: PB369P1D1
; INFORMATION FOR SEQ ID NO: 218:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 627 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 218:
US-10-206-576-218

Query Match      100.0%; Score 3225; DB 2; Length 627;
Best Local Similarity 100.0%; Pred. No. 1.5e-248;
Matches 627; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 MKQLKKWYTVSTLLILPLFTSVLGTTTAPAEENGESAQLVIHKKMTDLPDPLIONSG 60
        1 MKQLKKWYTVSTLLILPLFTSVLGTTTAPAEENGESAQLVIHKKMTDLPDPLIONSG 60
Db      61 KEMSEFDKYGGLADYTFISIYNTNFEYEQRAAGASVDAKQAVOSLTTPKPAQGTDDN 120
QY      61 KEMSEFDKYGGLADYTFISIYNTNFEYEQRAAGASVDAKQAVOSLTTPKPAQGTDDN 120
        61 KEMSEFDKYGGLADYTFISIYNTNFEYEQRAAGASVDAKQAVOSLTTPKPAQGTDDN 120
Db      61 KEMSEFDKYGGLADYTFISIYNTNFEYEQRAAGASVDAKQAVOSLTTPKPAQGTDDN 120
QY      61 KEMSEFDKYGGLADYTFISIYNTNFEYEQRAAGASVDAKQAVOSLTTPKPAQGTDDN 120
        61 KEMSEFDKYGGLADYTFISIYNTNFEYEQRAAGASVDAKQAVOSLTTPKPAQGTDDN 120
Db      121 GNVTVQLPKQNGKDAVYTIKEEPKEGVAAATNMVAPFVYEMIKQTDSYKGTBEELAV 180
QY      121 GNVTVQLPKQNGKDAVYTIKEEPKEGVAAATNMVAPFVYEMIKQTDSYKGTBEELAV 180
        121 GNVTVQLPKQNGKDAVYTIKEEPKEGVAAATNMVAPFVYEMIKQTDSYKGTBEELAV 180
Db      121 GNVTVQLPKQNGKDAVYTIKEEPKEGVAAATNMVAPFVYEMIKQTDSYKGTBEELAV 180
QY      121 GNVTVQLPKQNGKDAVYTIKEEPKEGVAAATNMVAPFVYEMIKQTDSYKGTBEELAV 180
        121 GNVTVQLPKQNGKDAVYTIKEEPKEGVAAATNMVAPFVYEMIKQTDSYKGTBEELAV 180
Db      181 VHIYKPNVANDGSLHVKKVGTAEENGGLNGAEFVLSKESGPGTVKXIQGVNDGLYTWTT 240
QY      181 VHIYKPNVANDGSLHVKKVGTAEENGGLNGAEFVLSKESGPGTVKXIQGVNDGLYTWTT 240
        181 VHIYKPNVANDGSLHVKKVGTAEENGGLNGAEFVLSKESGPGTVKXIQGVNDGLYTWTT 240
Db      181 VHIYKPNVANDGSLHVKKVGTAEENGGLNGAEFVLSKESGPGTVKXIQGVNDGLYTWTT 240
QY      181 VHIYKPNVANDGSLHVKKVGTAEENGGLNGAEFVLSKESGPGTVKXIQGVNDGLYTWTT 240
        181 VHIYKPNVANDGSLHVKKVGTAEENGGLNGAEFVLSKESGPGTVKXIQGVNDGLYTWTT 240
Db      241 DKEQAKRFITGKSYEIGENDFTEAENGTELTVKULEVGSYLLEEVKAPNNALLIENQTK 300
QY      241 DKEQAKRFITGKSYEIGENDFTEAENGTELTVKULEVGSYLLEEVKAPNNALLIENQTK 300
        241 DKEQAKRFITGKSYEIGENDFTEAENGTELTVKULEVGSYLLEEVKAPNNALLIENQTK 300
Db      241 DKEQAKRFITGKSYEIGENDFTEAENGTELTVKULEVGSYLLEEVKAPNNALLIENQTK 300
QY      241 DKEQAKRFITGKSYEIGENDFTEAENGTELTVKULEVGSYLLEEVKAPNNALLIENQTK 300
        241 DKEQAKRFITGKSYEIGENDFTEAENGTELTVKULEVGSYLLEEVKAPNNALLIENQTK 300
Db      301 TPFTIEANNQTPVEKTVKNDTSKVDKTPPSLDGKQVAIGEKKYQISVNIPLGIADKED 360
QY      301 TPFTIEANNQTPVEKTVKNDTSKVDKTPPSLDGKQVAIGEKKYQISVNIPLGIADKED 360
        301 TPFTIEANNQTPVEKTVKNDTSKVDKTPPSLDGKQVAIGEKKYQISVNIPLGIADKED 360
Db      301 TPFTIEANNQTPVEKTVKNDTSKVDKTPPSLDGKQVAIGEKKYQISVNIPLGIADKED 360
QY      301 TPFTIEANNQTPVEKTVKNDTSKVDKTPPSLDGKQVAIGEKKYQISVNIPLGIADKED 360
        301 TPFTIEANNQTPVEKTVKNDTSKVDKTPPSLDGKQVAIGEKKYQISVNIPLGIADKED 360
Db      361 ANKYKFNLVDRHDAALTFDNTVSGEYAYALYDGTVIAPENYQVTEQANGFTVAVNPAY 420
QY      361 ANKYKFNLVDRHDAALTFDNTVSGEYAYALYDGTVIAPENYQVTEQANGFTVAVNPAY 420
        361 ANKYKFNLVDRHDAALTFDNTVSGEYAYALYDGTVIAPENYQVTEQANGFTVAVNPAY 420
Db      361 ANKYKFNLVDRHDAALTFDNTVSGEYAYALYDGTVIAPENYQVTEQANGFTVAVNPAY 420
QY      361 ANKYKFNLVDRHDAALTFDNTVSGEYAYALYDGTVIAPENYQVTEQANGFTVAVNPAY 420
        361 ANKYKFNLVDRHDAALTFDNTVSGEYAYALYDGTVIAPENYQVTEQANGFTVAVNPAY 420
Db      421 IPTLPGGTLKRVYFMHINEKADPTKGFKNBANVDNGHTDDQTPPTVEVYVVGKGFIRKD 480
QY      421 IPTLPGGTLKRVYFMHINEKADPTKGFKNBANVDNGHTDDQTPPTVEVYVVGKGFIRKD 480
        421 IPTLPGGTLKRVYFMHINEKADPTKGFKNBANVDNGHTDDQTPPTVEVYVVGKGFIRKD 480
Db      421 IPTLPGGTLKRVYFMHINEKADPTKGFKNBANVDNGHTDDQTPPTVEVYVVGKGFIRKD 480
QY      421 IPTLPGGTLKRVYFMHINEKADPTKGFKNBANVDNGHTDDQTPPTVEVYVVGKGFIRKD 480
        421 IPTLPGGTLKRVYFMHINEKADPTKGFKNBANVDNGHTDDQTPPTVEVYVVGKGFIRKD 480
Db      481 GGVTAOTALAGASFVVRDQNSDTANYLKI DETTKAATVTKAABATFTTTADGLVDING 540
QY      481 GGVTAOTALAGASFVVRDQNSDTANYLKI DETTKAATVTKAABATFTTTADGLVDING 540
        481 GGVTAOTALAGASFVVRDQNSDTANYLKI DETTKAATVTKAABATFTTTADGLVDING 540
Db      481 GGVTAOTALAGASFVVRDQNSDTANYLKI DETTKAATVTKAABATFTTTADGLVDING 540
QY      481 GGVTAOTALAGASFVVRDQNSDTANYLKI DETTKAATVTKAABATFTTTADGLVDING 540
        481 GGVTAOTALAGASFVVRDQNSDTANYLKI DETTKAATVTKAABATFTTTADGLVDING 540
Db      541 LKGYTYLEETVAPDDYVLLTNRIEFVNEOSYGTTEMLVSPKYPNKHGTLPTSGGKG 600
QY      541 LKGYTYLEETVAPDDYVLLTNRIEFVNEOSYGTTEMLVSPKYPNKHGTLPTSGGKG 600
        541 LKGYTYLEETVAPDDYVLLTNRIEFVNEOSYGTTEMLVSPKYPNKHGTLPTSGGKG 600
Db      541 LKGYTYLEETVAPDDYVLLTNRIEFVNEOSYGTTEMLVSPKYPNKHGTLPTSGGKG 600
QY      541 LKGYTYLEETVAPDDYVLLTNRIEFVNEOSYGTTEMLVSPKYPNKHGTLPTSGGKG 600
        541 LKGYTYLEETVAPDDYVLLTNRIEFVNEOSYGTTEMLVSPKYPNKHGTLPTSGGKG 600
Db      601 IYVYLSGAVLLLIAGVYFARRRKENA 627
QY      601 IYVYLSGAVLLLIAGVYFARRRKENA 627
        601 IYVYLSGAVLLLIAGVYFARRRKENA 627
Db      601 IYVYLSGAVLLLIAGVYFARRRKENA 627

RESULT 3
US-09-134-000C-6124
; Sequence 6124, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
```

;; TITLE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS  
;; FILE REFERENCE: 032796-032  
;; CURRENT APPLICATION NUMBER: US/09/134,000C  
;; CURRENT FILING DATE: 1998-08-13  
;; PRIOR APPLICATION NUMBER: US 60/055,778  
;; PRIOR FILING DATE: 1997-08-15  
;; NUMBER OF SEQ ID NOS: 6812  
;; SOFTWARE: PatentIn version 3.1  
;; SEQ ID NO: 6124  
;; LENGTH: 659  
;; TYPE: PRT  
;; ORGANISM: Enterococcus faecalis  
US-09-134-000C-6124

Query Match 100.0%; Score 3225; DB 2; Length 659;  
Best Local Similarity 100.0%; Pred. No. 1.7e-248;  
Matches 627; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MQOLKKWVTASTLLILPLFTSVIGTTTAPAEENGESQVLVHKQMTDLPPLIQNSG 60  
DB 33 MQOLKKWVTASTLLILPLFTSVIGTTTAPAEENGESQVLVHKQMTDLPPLIQNSG 92  
QY 61 KEMSEFDKXQGLADVTFSIYVNTNEFEQRAAGASVDAKQAVQSLTPGKPVAAQTTDAN 120  
DB 93 KEMSEFDKXQGLADVTFSIYVNTNEFEQRAAGASVDAKQAVQSLTPGKPVAAQTTDAN 152  
QY 121 GNVTVQLPKQNGKQAVTIKEEPKEGVVAATMVAAPVYEMIKTDSYKGTBELAV 180  
DB 153 GNVTVQLPKQNGKQAVTIKEEPKEGVVAATMVAAPVYEMIKTDSYKGTBELAV 212  
QY 181 VHIYPPKNVANDGSLHVKVGTAEENGELNGAEFVLSKSGSPGTVCYIQGVKQGLYTTWT 240  
DB 213 VHIYPPKNVANDGSLHVKVGTAEENGELNGAEFVLSKSGSPGTVCYIQGVKQGLYTTWT 272  
QY 241 DKEQAKRFITGKSYEIGENDPTEAENGTELTIVKNLEVGSIYLEEVKAPNNAELIENQTK 300  
DB 273 DKEQAKRFITGKSYEIGENDPTEAENGTELTIVKNLEVGSIYLEEVKAPNNAELIENQTK 332  
QY 301 TPTTIEANNQTPVEKTVKNDTSKVDKTTPTSLDGKVAIGEKIKYQISVNIPLGIADKEGD 360  
DB 333 TPTTIEANNQTPVEKTVKNDTSKVDKTTPTSLDGKVAIGEKIKYQISVNIPLGIADKEGD 392  
QY 361 ANKYKFNLVGDHALTFPDNTSGEYAYALYDGDVIAPENYQVTEQANGFTVAANPAY 420  
DB 393 ANKYKFNLVGDHALTFPDNTSGEYAYALYDGDVIAPENYQVTEQANGFTVAANPAY 452  
QY 421 IPTLTPGGTLKFVYFMHLEKADPTKGFQKNEANVNDGHTDQTPPTVEVVTGSKREIKYD 480  
DB 453 IPTLTPGGTLKFVYFMHLEKADPTKGFQKNEANVNDGHTDQTPPTVEVVTGSKREIKYD 512  
QY 481 GCVTATQALAGASFYVRDQNSDTANYLKTIDETTKAATVYKTAQEAATFTTTADGLVDITG 540  
DB 513 GCVTATQALAGASFYVRDQNSDTANYLKTIDETTKAATVYKTAQEAATFTTTADGLVDITG 572  
QY 541 LKGYTYLEETVAPDDVYVLTNRIFVNVNOSYGTETNVPSPKVPNGKGLTLPSTGSKG 600  
DB 573 LKGYTYLEETVAPDDVYVLTNRIFVNVNOSYGTETNVPSPKVPNGKGLTLPSTGSKG 632  
QY 601 IYVYLGSGAVLLLIAGVYFARRRKENA 627  
DB 633 IYVYLGSGAVLLLIAGVYFARRRKENA 659

RESULT 4  
US-09-071-035-220

;; Sequence 220, Application US/09071035  
;; Patent No. 6448043  
;; GENERAL INFORMATION:  
;; APPLICANT: Gail H. Choi  
;; TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides  
;; NUMBER OF SEQUENCES: 496  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Human Genome Sciences, Inc.

;; STREET: 9410 Key West Avenue  
;; CITY: Rockville  
;; STATE: Maryland  
;; COUNTRY: USA  
;; ZIP: 20850  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Diskette, 3.50 inch, 1.4MB storage  
;; COMPUTER: HP Vectra 486/33  
;; OPERATING SYSTEM: MSDOS version 6.2  
;; SOFTWARE: ASCII Text  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/09/071,035  
;; FILING DATE:  
;; CLASSIFICATION:  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER:  
;; FILING DATE:  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: A. Anders Brookes  
;; REGISTRATION NUMBER: 36,373  
;; REFERENCE/DOCKET NUMBER: PB369P2  
;; TELEPHONE: (301) 309-8504  
;; TELEFAX: (301) 309-8512  
;; INFORMATION FOR SEQ ID NO: 220:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 560 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: protein  
US-09-071-035-220

Query Match 89.7%; Score 2892; DB 2; Length 560;  
Best Local Similarity 100.0%; Pred. No. 4.8e-222;  
Matches 560; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 BENGSAQVLVHKQMTDLPPLIQNSGKEMSEFDKXQGLADVTFSIYVNTNEFEQRA 92  
DB 1 BENGSAQVLVHKQMTDLPPLIQNSGKEMSEFDKXQGLADVTFSIYVNTNEFEQRA 60  
QY 93 GASVDAKQAVQSLTPGKPVAAQTTDANGNVTVQLPKQNGKQAVTIKEEPKEGVAAAT 152  
DB 61 GASVDAKQAVQSLTPGKPVAAQTTDANGNVTVQLPKQNGKQAVTIKEEPKEGVAAAT 120  
QY 153 NNVAAPVYEMIKTDSYKGTBELAVVHIYPPKNVANDGSLHVKVGTAEENGELNGAE 212  
DB 121 NNVAAPVYEMIKTDSYKGTBELAVVHIYPPKNVANDGSLHVKVGTAEENGELNGAE 180  
QY 213 FVLSKSGSPGTVCYIQGVKQGLYTTWTDEQAKRFITGKSYEIGENDPTEAENGTELT 272  
DB 181 FVLSKSGSPGTVCYIQGVKQGLYTTWTDEQAKRFITGKSYEIGENDPTEAENGTELT 240  
QY 273 VKNLEVGSIYLEEVKAPNNAELIENQTKPTTIEANNQTPVEKTVKNDTSKVDKTTPSLD 332  
DB 241 VKNLEVGSIYLEEVKAPNNAELIENQTKPTTIEANNQTPVEKTVKNDTSKVDKTTPSLD 300  
QY 333 GQDVAIGEKIKYQISVNIPLGIADYEGDANKYKFNLVGDHALTFPDNTSGEYAYALY 392  
DB 301 GQDVAIGEKIKYQISVNIPLGIADYEGDANKYKFNLVGDHALTFPDNTSGEYAYALY 360  
QY 393 DGDVIAPENYQVTEQANGFTVAANPAYIPTLTPGGTLKFVYFMHLEKADPTKGFQKNEA 452  
DB 361 DGDVIAPENYQVTEQANGFTVAANPAYIPTLTPGGTLKFVYFMHLEKADPTKGFQKNEA 420  
QY 453 NVNDGHTDQTPPTVEVVTGSKREIKYDGDVATATQALAGASFYVRDQNSDTANYLKTIDET 512  
DB 421 NVNDGHTDQTPPTVEVVTGSKREIKYDGDVATATQALAGASFYVRDQNSDTANYLKTIDET 480  
QY 513 TQAATVYKTAQEAATFTTTADGLVDITGLKGYTYLEETVAPDDVYVLTNRIFVNVNOS 572  
DB 481 TQAATVYKTAQEAATFTTTADGLVDITGLKGYTYLEETVAPDDVYVLTNRIFVNVNOS 540

QY 573 YGTTENLVSPKVPKHKGT 592  
|||  
Db 541 YGTTENLVSPKVPKHKGT 560

RESULT 5  
US-10-206-576-220  
; Sequence 220, Application US/10206576  
; Patent No. 6913907  
; GENERAL INFORMATION:  
; APPLICANT: Choi et al.  
; TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides  
; NUMBER OF SEQUENCES: 497  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Human Genome Sciences, Inc.  
; STREET: 9410 Key West Avenue  
; CITY: Rockville  
; STATE: Maryland  
; COUNTRY: USA  
; ZIP: 20850  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: CD-R  
; COMPUTER: Dell Latitude  
; OPERATING SYSTEM: Windows 98  
; SOFTWARE: ASCII Text  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/10/206,576  
; FILING DATE: 29-Jul-2002  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 09/071,035  
; FILING DATE: 1998-05-04  
; APPLICATION NUMBER: US 60/046,655  
; FILING DATE: 1997-05-16  
; APPLICATION NUMBER: US 60/044,031  
; FILING DATE: 1997-05-06  
; APPLICATION NUMBER: US 60/066,009  
; FILING DATE: 1997-11-14  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Hyman, Mark J.  
; REGISTRATION NUMBER: 46,789  
; REFERENCE/DOCKET NUMBER: PB369P1D1  
; INFORMATION FOR SEQ ID NO: 220:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 560 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; SEQUENCE DESCRIPTION: SEQ ID NO: 220:  
US-10-206-576-220

Query Match 89.7%; Score 2892; DB 2; Length 560;  
Best Local Similarity 100.0%; Pred. No. 4.8e-222;  
Matches 560; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 BENGSAOLVHKHKKMTDLPDLIONSGKEMSEPDKYQGLAVTTSIVNTEFEYORAA 92  
|||  
Db 1 BENGSAOLVHKHKKMTDLPDLIONSGKEMSEPDKYQGLAVTTSIVNTEFEYORAA 60

QY 93 GASVDAAKQAVSLTPGKFAQGTDDANGNVTVOLPKKONGDAVYTIKEBPKEGVAAAT 152  
|||  
Db 61 GASVDAAKQAVSLTPGKFAQGTDDANGNVTVOLPKKONGDAVYTIKEBPKEGVAAAT 120

QY 153 NNVVAFPVYEMTKQDGSYKYGTEELAVVHIIPKVVAVDGSLLHYKKVGTAEENGANGAE 212  
|||  
Db 121 NNVVAFPVYEMTKQDGSYKYGTEELAVVHIIPKVVAVDGSLLHYKKVGTAEENGANGAE 180

QY 213 FVTSSESGSPGVTKIYQVKGDLVYTTTDEKQAKRPTIGSKYEIGENDTEAENGTELT 272  
|||  
Db 181 FVTSSESGSPGVTKIYQVKGDLVYTTTDEKQAKRPTIGSKYEIGENDTEAENGTELT 240

QY 273 VKNLEVGSYLLEBVKAPNNABELIENQTKPTIEANNQTPVEKTVKNQTSKYDKTTPSLD 332  
|||

Db 241 VKNLEVGSYLLEBVKAPNNABELIENQTKPTIEANNQTPVEKTVKNQTSKYDKTTPSLD 300  
|||

QY 333 GKDVAIIGKIKYQISVNIPLGIADKEGDANKYVKNLVKDKDAALTENNUTSGEYAVLY 392  
|||  
Db 301 GKDVAIIGKIKYQISVNIPLGIADKEGDANKYVKNLVKDKDAALTENNUTSGEYAVLY 360

QY 393 DDDTVIAPENTQVEQANGFTVAVNPAYIPTLTPGTLKFFVFMHANEKADPTKGFKNBA 452  
|||  
Db 361 DDDTVIAPENTQVEQANGFTVAVNPAYIPTLTPGTLKFFVFMHANEKADPTKGFKNBA 420

QY 453 NVNDGHTDDQTPPVYEVVTVGKRFIKVDGVTATQALAGASFVVDQNSDTANYIKIDET 512  
|||  
Db 421 NVNDGHTDDQTPPVYEVVTVGKRFIKVDGVTATQALAGASFVVDQNSDTANYIKIDET 480

QY 513 TKAATWVTKABATFTTTADGLVDITGLKGYTYLEETVAPDDVLLTNRIEFVNBOS 572  
|||  
Db 481 TKAATWVTKABATFTTTADGLVDITGLKGYTYLEETVAPDDVLLTNRIEFVNBOS 540

QY 573 YGTTENLVSPKVPKHKGT 592  
|||  
Db 541 YGTTENLVSPKVPKHKGT 560

RESULT 6  
US-09-107-532A-7252  
; Sequence 7252, Application US/09107532A  
; Patent No. 6583275  
; GENERAL INFORMATION:  
; APPLICANT: Lynn A Doucette-Stamm and David Bush  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS  
; NUMBER OF SEQUENCES: 7310  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION  
; STREET: 100 Beaver Street  
; CITY: Waltham  
; STATE: Massachusetts  
; COUNTRY: USA  
; ZIP: 02354  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: CD-ROM ISO9660  
; COMPUTER: PC  
; OPERATING SYSTEM: <Unknown>  
; SOFTWARE: ASCII  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/107,532A  
; FILING DATE: 30-Jun-1998  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 60/085,598  
; FILING DATE: 14 May 1998  
; APPLICATION NUMBER: 60/051571  
; FILING DATE: July 2, 1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Ariniello, Pamela Deneke  
; REGISTRATION NUMBER: 40,489  
; REFERENCE/DOCKET NUMBER: GTC-012  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (781)893-5007  
; TELEFAX: (781)893-8277  
; INFORMATION FOR SEQ ID NO: 7252:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 664 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; HYPOTHEICAL: YES  
; ORIGINAL SOURCE:  
; ORGANISM: Enterococcus faecium  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (B) LOCATION 1...664  
; SEQUENCE DESCRIPTION: SEQ ID NO: 7252:  
;

US-09-107-532A-7252

Query Match 75.8%; Score 2446; DB 2; Length 664;

Best Local Similarity 74.2%; Pred. No. 2,4e-186; Matches 462; Conservative 64; Mismatches 95; Indels 2; Gaps 1;

```
QY 1 MQLKQWVTVSTLILPLFTSVLGTTPAFAENGESQVLVHKKKMTDLPPLQNSG 60
DB 40 MKNHKKINMLGVFLILPLTNSFPAKKVFAET--AAOVILHKKMTDLPPLQNSG 97
QY 61 KEMSEBDYQGLADVTFSYNTNEFEYQRAAGASVDAKQAVSLTPGKPAVQGTDDAN 120
DB 98 KEMSEBDYQGLADISFSYNTNEFEYQRAAGASVDAKQAVSLTPGKPAVQGTDDAD 157
QY 121 GNVTVOLPKKQKQDAVTTIKKEPKGVAATMVAAPVYEMIKOTDGSYKKGTEBLAV 180
DB 158 GNVTVSLPKKQKQDAVTTIKKEPKGVAATMVAAPVYEMIKOTDGSYKKGTEBLDT 217
QY 181 VHIYKRVNVANGSLHVKVGTAEENGANGAFVLSKSGSPGTVYKIQGVQDGLYTWTT 240
DB 218 IHLVYKRVNVANGSLHVKVGTAEENGANGAFVLSKSGSPGTVYKIQGVQDGLYTWTT 277
QY 241 DKOAKRFITGKSYEIGENDPTEAENGTELTVKNLEVGSYLIEYKAPNNALLENQTK 300
DB 278 DGTAKAHFITHSYDIGNDDFAASIEKQOLIVNHLEVGKYNLEBKADPNMEMIKQTI 337
QY 301 TPTTEANNQTFVEKTVKNDTSKVDKTTPSLDGKQVAGIKIKYQISVNIPLGIADKEGD 360
DB 338 TPEELIANGQTFVEKTIKNDTSKVDKTTPOLNKGQVAGIKIQYEISVNIPLGIADKEGT 397
QY 361 AKRYKENVLDHDAALTPDNVTSGEYAVAYLDGQDVIAPENYQVTEANGFTVANPAY 420
DB 398 QNRKTYTFKIDITHDAALTPDNDSSTGYAVALYDGNKEIDVNSVTEQDGFVSDPENY 457
QY 421 IFTLPGGTLKFVYFMHNEKADPTKGFKNENAVNDGHTDDQTPPYEVVVGSKREIKYD 480
DB 458 ISLTPTGGLKFVYFMHNEKADPTKGFKNENAVNDGHTDDQTPPYEVVVGSKREIKYD 517
QY 481 GNVTAQALAGASFVVRDQNSDTANYLIKIDETTKAATWVTKAETPTTTADGLVDITG 540
DB 518 GNVTSQTLAGAFVVRDQSDPTAKYLSIDSTKAVSWSAKESATVFTTNSGLIDVTG 577
QY 541 LKYGTYLLEETVAPDDYVLLTNRIEFVNVNQSSTENLVSPKVNKHKGLPSTGGG 600
DB 578 LKYGTYLLEETVAPDEYVLLTNKVAFTIDQSYVTAGQLISPKIKHKKGLPSTGGG 637
QY 601 IYVYLGSAVLLLIAGVFPARR 623
DB 638 IYVYLGSAVLLLIAGVFPARRK 660
```

RESULT 7

```
US-09-134-000C-6123
; Sequence 6123, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; PRIOR FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6123
; LENGTH: 477
; TYPE: PRN
; ORGANISM: Enterococcus faecalis
US-09-134-000C-6123
Query Match 13.6%; Score 438.5; DB 2; Length 477;
```

Best Local Similarity 25.9%; Pred. No. 2,1e-26; Matches 166; Conservative 70; Mismatches 211; Indels 193; Gaps 22;

```
QY 8 WTVVSTLILILPLFTSVLGTTPAFAENGESQVLVHKKKMTDLPPLQNSGKEMSE 65
DB 7 WLSICWMLALFGFSQ-----QALAEASQASQVTLHKKLFPDGOQLPEOD-QNTGSEGL 60
QY 66 FDKYQGLADVTFSYNTNEFEYQRAAGASVDAKQ--AVQSLTPGKPAVQGTDDA---- 119
DB 61 LQNTREGLNDVTVQVNDVTPDYQLNSBEGKTQBAQQLAETGATNRKPLAEDTKTQIINGE 120
QY 120 NGNVTVOLPKK--QNGKDAVTTIKKEPKGVAATMVAAPVYEMIKOTDGSYKKGTE 176
DB 121 DGVNFSLSLXOSQQRDKAVLVFEABAEPVYEMIKSNLVVILPYQDPQSGS----- 171
QY 177 ELAVHITYKRVNVANGSLHVKVGTAEENGANGAFVLSKSGSPGTVYKIQGVQDGLY 236
DB 172 -LTHIHLVYKKN----- 181
QY 237 TWTTDEQAKRFITGKSYEIGENDPTEAENGTELTVKNLEVGSYLIEYKAPNNALLE 296
DB 182 -----EENAYDL----- 188
QY 297 NOTKPTTEANNQTFVEKTVKNDTSKVDKTTPSLDGKQVAGIKIKYQISVNIPLGIAD 356
DB 189 -----PLEKTVLDKQGFNQ-----GEHINQVLTQIIPANILG 222
QY 357 KEGDANKYKENVLDHDAALTPDNVTSGEYAVAYLDGQDVIAPENYQVTEANGFTVAN 416
DB 223 -----YQEFRLSDKADTLTL-----LPRESIEVKVAGKVTVT--GYTLTQHGHTPLDF 269
QY 417 NPAVITPLTPGGLTKFVYFMHNEKADPTKGFKNEN--VDNGHTDQTPPYEVVVGGR 475
DB 270 SIKDLQNPRA--NOTMVSYQMLEKTAEPDTAINNESQVLVDKHT--LTKRATVRIGAS 325
QY 476 FIKVDGTVATQALAGASFVVRDQNSDTANYLIKIDETTKAATWVTKAETPTTTADGL 535
DB 326 FVKVDSR--NAKTLTPAIVIVKQAGEYLN-----ETANGYKQKELAKKTTSNQAGE 379
QY 536 VDTGLKGYTYLLEETVAPDDYVLLTNRIEFVNVNQSSTENLVSPKVPN--XHKGT 592
DB 380 FSVKGLKQGYFLIEETISAPKGYLINTBIPFTVGKNSYATNGQRTAPLHVINKVSESG 439
QY 593 LPSTG-GKGIYVYLGSAVLLLIAGV-----YFPARRKE 625
DB 440 LPTMERSISWLTIAG---LLIIGNVVIMLFYQKQKE 475
```

RESULT 8

```
US-09-107-532A-6715
; Sequence 6715, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESS: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: PC
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532A
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
```

```

; APPLICATION NUMBER: 60/085,598
; FILING DATE: 14 May 1998
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 6715:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 660 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Enterococcus faecium
;
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...660
; SEQUENCE DESCRIPTION: SEQ ID NO: 6715:
US-09-107-532A-6715

Query Match      12.9%; Score 414.5; DB 2; Length 660;
Best Local Similarity 25.3%; Pred. No. 2.9e-24;
Matches 187; Conservative 102; Mismatches 247; Indels 203; Gaps 38;

QY      2 KOLKVVYVSTLLILPLFTSVLGTTFAPAE-----NGESAOLVIHKKMTDLPDP- 55
DB      6 RIVKAGMLVGMILGIIAGCFQGY---RAYAEDVYOKTPEKXNITVH-KLMTYDQGTQLN 60
QY      56 -----IUNGKEMSEF---DKYQ--GLADVTFSTYVNTNFEYEQRAA----- 92
DB      61 VDVVDGKNDGYTHDQYPTGVTKYKNRADYGDVEFTLGNITDQVLPREDSDLTNAKVDIYK 120
QY      93 -----GASVDAKQAVQSLTPGKPVAAQGTTDANGNVTV--QLPKONGCDAYITIKE-EP 144
DB      121 DVEDKGSNSEYVKNATNKIT-----SAVDERNGEITPADQAPAYVNSKGVYVYVESKS 173
QY      145 KEGVVA--ATNNVVAFPVYEMIKOTDGSYKYGEBELAVVHIYKPVNVANDGSLHYKV-- 200
DB      174 AAGLVYQRAKPVNVIAFM-----TDNT---SSGFLKDIHLTPKAIIVSR-LSFELTKGSD 223
QY      201 -GTAENE--GLNGAEFVISKSESGSPGVKYIQGVXDGLYTWTTDKEQAKRFTGKSVEIG 257
DB      224 DGTAGSKOTPLKGAKEFLYKGE-----PGKGTKLGG 253
QY      258 ENDFTAEANGELVVKULEVGSYILLEEKAPNNAELINQTKTPT-----IBA 307
DB      254 D---LVSDDOQKILATDLTLGKYFVEVP---SEVVGSDKEPFAADQYLLGADARADA 305
QY      308 NNQTFVEKTVKNDTSK-----VDKTTPSLDGKD--VAIGEKIKYOISVNIPLG 353
DB      306 HNKLTFLITNDGVTSDLKASVYVYKAPVLDKTYVNTGTGGEHSFQIGDAVNYGTHIHPD 365
QY      354 IADKEGDANKYVKNLVKHDNALTFDNTVSGEYAYALYD--GDT----- 396
DB      366 IA--GGAD-----GITVNGVKSETSPYSVFKWMDTLAGQGLSYAAKAIKIV 409
QY      397 -----VIAPEV--YQVTEQANGFTV--AVNPAYITPLTP---GGLTKVYVFMHLNEKAD 443
DB      410 TNKDGSVYLKENTDVKIÖNSENGFVIDPVLNNGOVSDTVASLHGQDLQWTYMYMYVND5AA 469
QY      444 PTGKGFKEAN--VDNGHTDQ-----TPPTVEVNTVGGKRFIKVDGDTAT-----QALAGASF 494
DB      470 VANPLTNSVDFTYNNPNPQÖEHHEKTKADVVITYGAKFLKVDSGFLGTGIKATPLBSAEF 529
QY      495 VVRDONSPTANYLKIDET---TKAATVWKTAEATPTTTTADGLVDITGLKXGYTYLEE 550
DB      530 AAK--NABGKYGGVLVDTDKNGVKEAVWVDVANAILKSDKEGHEIIGLGEGEYSLEE 587
```

```

QY      551 TVAPDDVYLLTNRIFVYVNEQSYGTENTVSPKYPNKHKGLTLPSTGKGIVY----- 604
DB      588 TVAPENYQVLTKEISFKVDKQSY-KEENRT--IKNNQKASVPMTGSGNGQTYVILISCL 643
QY      605 -LGSAGVLLIAGVYFAR 622
DB      644 LLGAGA---LSAVVYFFKKK 659

RESULT 9
US-09-107-532A-5331
; Sequence 5331, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESSER: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: PC
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532A
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/085,598
; FILING DATE: 14 May 1998
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 5331:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 508 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Enterococcus faecium
;
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...508
; SEQUENCE DESCRIPTION: SEQ ID NO: 5331:
US-09-107-532A-5331

Query Match      12.2%; Score 395; DB 2; Length 508;
Best Local Similarity 24.4%; Pred. No. 6.9e-23;
Matches 156; Conservative 79; Mismatches 200; Indels 204; Gaps 26;

QY      4 LKV--MYTVST--LLILPLFTSVLGTTFAPAEENGESAOVLVHKKMTD--LP--DPLIQ 57
DB      36 MKGLGMLSMCLFLILFKRAFTQVA-----TETETEMVOITLHKLLFPNGQLPKXHP--- 86
QY      58 NSGKEMSEFDKYQGLADVTFSTYVNTNFEYEQRAAGASVDAKQ--AVQSLTPGKPVAAQ 115
DB      87 NDQGEKALLQTVYRGVNGVTFYQVYDVTDSEFHLREKGTVEEQAQAEIADKAGASSGMFTAEA 146
```





APPLICATION NUMBER: US/10/206,576  
FILING DATE: 29-Jul-2002  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 09/071,035  
FILING DATE: 1998-05-04  
APPLICATION NUMBER: US 60/046,655  
FILING DATE: 1997-05-16  
APPLICATION NUMBER: US 60/044,031  
FILING DATE: 1997-05-06  
APPLICATION NUMBER: US 60/066,009  
FILING DATE: 1997-11-14  
ATTORNEY/AGENT INFORMATION:  
NAME: Hyman, Mark J.  
REGISTRATION NUMBER: 46,789  
REFERENCE/DOCKET NUMBER: PB369P1D1  
INFORMATION FOR SEQ ID NO: 274:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 430 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 274:  
US-10-206-576-274

Query Match 10.8%; Score 347; DB 2; Length 430;  
Best Local Similarity 24.0%; Pred. No. 3,6e-19; Indels 186; Gaps 19;  
Matches 142; Conservative 64; Mismatches 200;  
QY 8 WYVSTLLILPLFTSVLGTTPAFENESAOVYHKKMTD--LPPELLQNSKENSE 65  
DB 6 WLSICVMLLALFGFSG-----QALAEASQASVQVTLHLKLPFGDGLPEQO-QNTGEBGL 59  
QY 66 FDKYGLADVFTSYNVTNEFYEQRAAGSVDAKO--AVQSLTEGKPYAGSTTA---- 119  
DB 60 LQNYGLNDVTVQVYDVTDPFYQLRSEKTVQEAQROLAETGATNRKPIAEDKTQTINGE 119  
QY 120 NGNVTVOLPKK---QNGKDAVYTIKEPREGVVAATNMVAPPVYEMIKQTGSKYGYE 176  
DB 120 DGVSFSLASKDSQQRDKAYLFVEAEAPFVVEKASNLVILPVDQPOGQS----- 170  
QY 177 ELAVVHIYKQNVVANDGSLHVKKVGTAEENGLNGAEFVLSKESGPGTVKYIQGVGDLY 236  
DB 171 -LTHIHLYPKN----- 180  
QY 237 TWTTTKEQAKRFTYKSYEIGENDFTEANGTGLTVKNLVEGSYLLEEVKAPNNAELIE 296  
DB 181 -----EENAYDL----- 187  
QY 297 NOTKPTTEANNQTPVEKTVKNDTSKVDKTPPSLDGKDAVAGEKIKYQISVINPIGLAD 356  
DB 188 -----PPELKTIVLDKQOGFNO-----GEHINYQLTQITANILG 221  
QY 357 KEGDANKYKFNLDKHDAAALTFDNTVSGEYAYALYDGPVIAPEHYQTEGANGFTYAV 416  
DB 222 -----YQEFRISDKADTTLT-----LPESIEVKVAGKTVTT--GYTLTQKHGFTLDF 268  
QY 417 NPAYIPTLPGTGLKFVFMHLNEKADPTKGFQNEAN--VDNGHTDDQTPPYEVVTTGGR 475  
DB 269 SIKDQONFA-NQTMVTSYQMRLEKTAEPDTAINNSEQVLTDKGT--LTKRAVTRGKS 324  
QY 476 FIKVGDVATQALAGASFVVDQNSDPTANYLKISETTKAATWVTKAEATFTTTADL 535  
DB 325 FVKVDSE-NAKTLTLEAVPIVKNQAGEYLN-----ETANGYMOKEKALAKFTTSNAGE 378  
QY 536 VDIGLKYGTYYLLEETVAPDDVYLLTNRIEFVNNQSYGTTENLVSEKVPN 587  
DB 379 FSVKXKKEPVLIGRN-----LCTKRLSSESNNRNSFYGKKFLCNERRTN 423

RESULT 12  
US-09-071-035-276

Sequence 276, Application US/09071035  
Patent No. 6448043  
GENERAL INFORMATION:  
APPLICANT: Gil H. Choi  
TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides  
NUMBER OF SEQUENCES: 496  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Human Genome Sciences, Inc.  
STREET: 9410 Key West Avenue  
CITY: Rockville  
STATE: Maryland  
COUNTRY: USA  
ZIP: 20850  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.44mb storage  
COMPUTER: HP Vectra 486/33  
OPERATING SYSTEM: MSDOS version 6.2  
SOFTWARE: ASCII text  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/071,035  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: A. Anders Brookes  
REGISTRATION NUMBER: PB369P2  
REFERENCE/DOCKET NUMBER: PB369P2  
TELEPHONE: (301) 309-8512  
TELEFAX: (301) 309-8504  
INFORMATION FOR SEQ ID NO: 276:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 384 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-071-035-276

Query Match 9.6%; Score 311; DB 2; Length 384;  
Best Local Similarity 23.5%; Pred. No. 2,2e-16;  
Matches 127; Conservative 56; Mismatches 180; Indels 178; Gaps 16;  
QY 57 QNSGKENSEPDKYGLADVFTSYNVTNEFYEQRAAGSVDAKO--AVQSLTEGKPYAQ 114  
DB 7 QNTGEBGLNQNYRGLNDVTVQVYDVTDPFYQLRSEKTVQEAQROLAETGATNRKPIAE 66  
QY 115 GTTDA---NGNVTVOLPKK---QNGKDAVYTIKEPREGVVAATNMVAPPVYEMIKQT 167  
DB 67 DKTQRTINEDGVVSFSLASKDSQQRDKAYLFVEAEAPFVVEKASNLVILPVDQPOGQS 126  
QY 168 DGSYKTYGTEBLAVVHIYKQNVVANDGSLHVKKVGTAEENGLNGAEFVLSKESGPGTVKY 227  
DB 127 -----LTHIHLYPKN----- 136  
QY 228 IQGVGDLYTWTTDKEQAKRFTYKSYEIGENDFTEANGTGLTVKNLVEGSYLLEEVK 287  
DB 137 -----EENAYDL----- 143  
QY 288 APNNAELIENQTKPTTEANNQTPVEKTVKNDTSKVDKTPPSLDGKDAVAGEKIKYQIS 347  
DB 144 -----PPELKTIVLDKQOGFNO-----GEHINYQLT 168  
QY 348 VNIPIGLADKEGDANKYKFNLDKHDAAALTFDNTVSGEYAYALYDGPVIAPEHYQTE 407  
DB 169 TQIPANILG-----YQEFRISDKADTTLT-----LPESIEVKVAGKTVTT--GYTLT 215  
QY 408 QANGFTVAVNPAYIPTLPGTGLKFVFMHLNEKADPTKGFQNEAN--VDNGHTDDQTPPT 466  
DB 216 QKHGFTLPSIKDQONFA-NQTMVTSYQMRLEKTAEPDTAINNSEQVLTDKGT--LTKR 271

```

QY 467 VEVNMGKRRPIKVDGVTANQALAGASFEVVRDQSDTANVLIKIDETTKAATWVTKREAT 526
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 272 ATVRGKGKSPVKDSE-NAKITLPFAVFNQNGEYLN-----ETANGYRQKEKALAK 325
QY 527 TETTTAGLVDTIGLKYGYLYLEETVAPDDVLLTNRIEFVNEQSQGTENTLVSPEKVP 586
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 326 KETSNAGSPSVGKXKRWPLLGRN-----LCTKRLSSNSNNSPYGKKFLCNEKTT 378
QY 587 N 587
    |
DB 379 N 379

RESULT 13
US-10-206-576-276
; Sequence 276, Application US/10206576
; Patent No. 6913907
GENERAL INFORMATION:
APPLICANT: Choi et al.
TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides
NUMBER OF SEQUENCES: 497
CORRESPONDENCE ADDRESSES:
ADDRESSER: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: CD-R
COMPUTER: Dell Latitude
OPERATING SYSTEM: Windows 98
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/206,576
FILING DATE: 29-Jul-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 09/071,035
FILING DATE: 1998-05-04
APPLICATION NUMBER: US 60/046,655
FILING DATE: 1997-05-16
APPLICATION NUMBER: US 60/044,031
FILING DATE: 1997-05-06
APPLICATION NUMBER: US 60/066,009
FILING DATE: 1997-11-14
ATTORNEY/AGENT INFORMATION:
NAME: Hyman, Mark J.
REGISTRATION NUMBER: 46,789
REFERENCE/DOCKET NUMBER: PB369PDI1
INFORMATION FOR SEQ ID NO: 276:
SEQUENCE CHARACTERISTICS:
LENGTH: 384 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 276:
US-10-206-576-276

Query Match 9.6%; Score 311, DB 2, Length 384;
Best Local Similarity 23.5%; Pired. No. 2,2e-16;
Matches 127; Conservative 56; Mismatches 180; Indels 178; Gaps 16;

QY 57 ONSGKEMSFDDKQGLADVTFSYNTNEPEQRAAGASVDAKO--AVGSLTFGKXPVQ 114
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 7 QNTGEGTILLONRGANDVTYGVYDVPYQLSSEKTVQEOARQLAETGATRKRPALAE 66
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 115 GTTDA---NGNTVOLPKK---QNGKDAVYTIKEEPKEGVAAITNNVVAEPVYEMIKQT 167
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 67 DDTGRTINGSDGVVFSLSKDSQGRDKAYLFVEAEAPVAVKESKSNLVLLVPQDPQGS 126
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 168 DGSYKKGTEELAVVHIYPKKNVANDGSLHYKKQGTAEENEGLNGAEFYISKSEGSFGTYKY 227
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

|    |     |                                                              |                             |     |
|----|-----|--------------------------------------------------------------|-----------------------------|-----|
| Db | 127 | -----                                                        | :                           | 136 |
| Qy | 228 | IQGVKDGLYTTWDBEOAKRFITGKSYBIGENDPTEAENGTEGLTVKCNLEVGSYILEBYK | -----LTHILYPRKN             | 287 |
| Db | 137 | -----                                                        | :                           | 143 |
| Qy | 288 | APNNAELIENQTKTPPTIEANNQTEPVKNDTSKVDKTPSSLDGKDAVIGKIKYQIS     | -----EENAYDL                | 347 |
| Db | 144 | -----                                                        | :                           | 168 |
| Qy | 348 | VNIPGLADKSGDANKYVKFNLYVDGHDAALFTPDNTSGEYALALYDQTVIAPENYQYE   | -----PLEKTVLDKQOGFN         | 407 |
| Db | 169 | TOIPNIIIG-----                                               | :                           | 215 |
| Qy | 408 | QANGPTAVNPAVYIPTLTPGGTLKVFVFMHLEKADPTGFKFNEAN-VDNGHITDDQTPPT | -----YQFRLSDRADTTLTL        | 466 |
| Db | 216 | QKHGTTLPSSIKDLONFA-NQTMWTSYQWRLEKTAEPDPAINNEGQVTDKHT--LTKR   | -----LPRSEIKVAKGVTT--GYTLTT | 271 |
| Qy | 467 | VEVWTGGRFLIKVDGVTATOLAGASFPVRQDSDNLTANYLKIDETTKATATWKTAEAT   | -----                       | 526 |
| Db | 272 | ATVRTRGGSFPKVBSE-NAKITLPEAVFIYKQQAEBYLN-----ETANGYRMOKEKALAK | -----                       | 325 |
| Qy | 527 | TFTTTADGLVDTIGLKIGTYIIEFTVAPDDVYLLTNRIEFVNVEDOSYGTENLVSPKVP  | -----                       | 586 |
| Db | 326 | KFTSNQAGEFVSVKGRWPLLGRN-----LCTKRLLSSBNRNSFYGKKFPLCNERTT     | -----                       | 378 |
| Qy | 587 | N 587                                                        | -----                       |     |
| Db | 379 | N 379                                                        | -----                       |     |

```

RESULT 14
US-09-769-787-127
; Sequence 127, Application US/09769787
; Patent No. 6936252
; GENERAL INFORMATION:
; APPLICANT: Microbial Technics Limited
; APPLICANT: Gilbert, Christophe FG
; APPLICANT: Hanabro, Phillip M
; TITLE OR INVENTION: Proteins
; FILE REFERENCE: PWC/P2129WO
; CURRENT APPLICATION NUMBER: US/09/769,787
; CURRENT FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: GB 9816337.1
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: US 60/125164
; PRIOR FILING DATE: 1999-03-19
; NUMBER OF SEQ ID NOS: 388
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 127
; LENGTH: 665
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-769-787-127

Query Match          9.2%; Score 296.5; DB 2; Length 665;
Best Local Similarity 23.0%; Pred. No. 7,66-15;
Matches 176; Conservative 84; Mismatches 259; Indels 247; Gaps 38

QY      1 MKQLKKVWYVTVSTLLILPFTSVLGTTFAEENGBSAQLVHKKQKTDLPDLLIONSG 60
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      1 MKSINKETLMLAALLLTA--SLSPSATVPA-AGTTTSTVHKLTALDDMDKIANE- 55
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      61 KENSEPDKYGS-----LADYTESLYNTNTEYEG--BAAGSVAARQAQVOSL 106
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      56 ---LETGNYAGNKKGVLPAANKEIAGVWFVWNTNNELIIDENQGLGVNIDPOTFKJSGA 112
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      107 TPGRVVAQGTTDANGNV--TVQLPKONGKDAVYTIKEEPKEGVAATNNVVAEPVYEMI 164
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      113 MPATVAMKK-LTEADGAFNTANLP-----AAKYKIYEIH 145
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      165 KQT-----DSGYKGTTELAV-----VHIYPRV-----VANDGSLH 196
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

Db 146 SLSTVVGSDGATLTGSKAVPIEILPLNDVVAHAYPKKTEAKPKIDKPKKANPDPBR 205  
QY 197 VKKGTAEENEGANGAEF-VISKSBSPGTVKTIQGVKGLYWTDTKQAKFPITGSKYE 255  
Db 206 VDKDTPVNHQVGDVVEYEIVTK--IPALANATA-----SDMTGLAF 249  
QY 256 IGENDFTEANGTGLTYKKN--LEVGSYLIEVK-----APNNAELINQTKT 301  
Db 250 -----NKGTAVYVDVALLEAGDVALTVAGFPLKLTDAGLAKKNDQNAEKTYKI 300  
QY 302 PFTTEANNQTPVEKTVKND-----TSKVDKTPSLDGDVAIGEKIKYQISV 348  
Db 301 TYSATLNDKRAIVEESNDVTFNNGNPDHGTPKPKNK--PNENG-DLTLTKTWDAIGA 357  
QY 349 NIPGLIADKEGDANKYKFNVLVDKIDALTFD--NTSSEYAYLYDGDVTYAPENYQVT 406  
Db 358 PIPAG-----AEATFDLVNAQTKGVQTV---TLTTDRN---T 389  
QY 407 EQANGFTVAVNPAYIPTLPGLTKFVYFMHLNEKADPTKGFKNENAVNDGHTDDQTPPT 466  
Db 390 VTVNLGDKTEKTFYERSIKGYSADYQELTTAGEIA--VKWKDE---NPKPLDPTPE- 442  
QY 467 VEVVTGGRKFIKY-DGDVATQALAGASFVVDONSDTA--NYL--KIDE-----T 512  
Db 443 -KVVTYGGKFFVKNKD-----NRLAGAEFVI--ANADNAGQYLARKADKVSQEKOLVVT 495  
QY 513 TK-----AATV--KTAKAATTF 528  
Db 496 TKDALDRAVAAVNALTAQOQTQOEKEKVDKQAAYNAVIAANNAFEWADNDONENVKL 555  
QY 529 TTTADGLVDITGLKGTYYLEETVAPDDVLLTNRLEFVNNEQSYGTTEENVL----- 580  
Db 556 VSDAGREBITLGLTYLEETKOPAGALALTSKQKEVATATYSATGQGIETAYAGSK 615  
QY 581 -SPEKVPNGHKGTLPSYTGKGIYVYLGSAVLLLIAGVYFARRRKE 625  
Db 616 DDATKVVNK-KITIPOTGGIGTIFAVAGAAIMGIAVYVAVYVKNND 660

RESULT 15  
US-09-107-532A-5384  
; Sequence 5384, Application US/09107532A  
; Patent No. 6583275  
; GENERAL INFORMATION:  
; APPLICANT: Lynn A Doucette-Stamm and David Bush  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS  
; NUMBER OF SEQUENCES: 7310  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION  
; STREET: 100 Beaver Street  
; CITY: Waltham  
; STATE: Massachusetts  
; COUNTRY: USA  
; ZIP: 02354  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: CD-ROM ISO9660  
; COMPUTER: PC  
; OPERATING SYSTEM: <Unknown>  
; SOFTWARE: ASCII  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/107,532A  
; FILING DATE: 30-Jun-1998  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 60/085,598  
; FILING DATE: 14 May 1998  
; APPLICATION NUMBER: 60/051571  
; FILING DATE: July 2, 1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Ariniello, Pamela Deneke  
; REGISTRATION NUMBER: 40,489  
; REFERENCE/DOCKET NUMBER: GTC-012

TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (781)893-5007  
; TELEFAX: (781)893-8277  
; INFORMATION FOR SEQ ID NO: 5384:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 341 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; HYPOTHEITICAL: YES  
; ORIGINAL SOURCE:  
; ORGANISM: Enterococcus faecium  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (B) LOCATION 1..341  
; SEQUENCE DESCRIPTION: SEQ ID NO: 5384:  
US-09-107-532A-5384  
Query Match 9.0%; Score 290.5; DB 2; Length 341;  
Best Local Similarity 27.1%; Pred. No. 8e-15;  
Matches 90; Conservative 58; Mismatches 125; Indels 59; Gaps 13;  
335 DVAIGEKIKYQISVNIPLGIADKEGDANKYK-FMLVDKIDALTFDNTSGEYAYLYD 393  
Db 23 DLERGKTASYTTIAPIDPYFIDSVLENGSAVINKYKITDTPVGLTY-----YD 70  
QY 394 -----GDVYIAP-ENYQVTEOANGFTVAV-----NPAYIPTL-----TPGTLKPYF 435  
Db 71 QEIVRAGETILTCKQDIIVEVNSGVYVITLITEENGVAKYD/LORLADAGGDLTIYNN 130  
QY 436 MHLNEKADPTKGFKNENAVNDGHTD-----DQTPTEVVTGGRFIVDGDVTATQALA 490  
Db 131 LKVTLEBADPFHNTAVIIEGRNDEPDYEEGVEPEKVTGGKFEKY--DASSSELK 188  
QY 491 GASFVVRQNSDTAY-----LKIDETTKATWKTAAEATFTTTADGLVDITGL 541  
Db 189 DARFEL--WNEDRSYALFYGESPLAYESGADRIEATSGQATEFVADGNGYFVQGL 246  
QY 542 KGTYYLEETVAPDDVLLTNR---IEFVNNEQSYGTTEENVL-----SPEKVPNGHKGTL 593  
Db 247 DYGTYQMETNAPEGVYLPTEGAATFEPIISGSINERIQIVGVNPPGERVPNNKRGSL 306  
QY 594 PSTGKGIYVYLGSAVLLLIAGVYFARRRKE 625  
Db 307 PATGNGGLAFLILGISIMI--GAYSWYRSK 336

Search completed: July 6, 2006, 02:38:16  
Job time : 52 secs

GenCore version 5.1.9  
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 5, 2006, 18:08:58 ; Search time 7250 Seconds  
(without alignments) 16591.048 Million cell updates/sec

Title: US-10-661-809A-12

Perfect score: 1881  
Sequence: 1 atgaagcaatcaaaaaaagt.....gacgcagaagaagaatgct 1881

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 6366136 seqs, 31973710525 residues

Total number of hits satisfying chosen parameters: 12732272

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

GenEmbl:  
1: gb\_env:\*  
2: gb\_pat:\*  
3: gb\_ph:\*  
4: gb\_pl:\*  
5: gb\_pr:\*  
6: gb\_ro:\*  
7: gb\_str:\*  
8: gb\_sy:\*  
9: gb\_un:\*  
10: gb\_vl:\*  
11: gb\_ov:\*  
12: gb\_hcg:\*  
13: gb\_in:\*  
14: gb\_cm:\*  
15: gb\_da:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score  | Query Match | Length | DB ID | Description                    |
|------------|--------|-------------|--------|-------|--------------------------------|
| 1          | 1881   | 100.0       | 1980   | 2     | AR396704 Sequence              |
| 2          | 1881   | 100.0       | 1983   | 2     | BD191807 Enterococ             |
| 3          | 1881   | 100.0       | 1983   | 2     | AR228552 Sequence              |
| 4          | 1881   | 100.0       | 1983   | 2     | AR692967 Sequence              |
| 5          | 1881   | 100.0       | 15614  | 2     | BD193451 Enterococ             |
| 6          | 1859   | 98.8        | 1875   | 2     | C0969453 Sequence              |
| 7          | 1859   | 98.8        | 110000 | 15    | AE016830_10 Continuation (11 o |
| 8          | 1687   | 89.7        | 1687   | 2     | BD191808 Enterococ             |
| 9          | 1687   | 89.7        | 1687   | 2     | AR228553 Sequence              |
| 10         | 1687   | 89.7        | 1687   | 2     | AR692968 Sequence              |
| 11         | 1014.6 | 53.9        | 1878   | 2     | C0969772 Sequence              |
| 12         | 1014.6 | 53.9        | 1995   | 2     | AR348987 Sequence              |
| 13         | 595    | 31.6        | 2199   | 2     | AX476900 Sequence              |
| 14         | 94.6   | 5.0         | 150038 | 12    | AC181263 Strongylo             |
| 15         | 94.6   | 5.0         | 204539 | 12    | AC178783 Strongylo             |
| 16         | 83.8   | 4.5         | 2069   | 2     | E10125 DNA encodin             |
| 17         | 83.8   | 4.5         | 3399   | 2     | E10126 DNA encodin             |
| 18         | 83.4   | 4.4         | 223469 | 12    | AC112451 Rattus no             |

|    |      |     |        |    |                                |
|----|------|-----|--------|----|--------------------------------|
| 19 | 80   | 4.3 | 1701   | 2  | CS075468 Sequence              |
| 20 | 80   | 4.3 | 1701   | 2  | CS082924 Sequence              |
| 21 | 80   | 4.3 | 253001 | 13 | AE014834 Plasmodu              |
| 22 | 79.2 | 4.2 | 198224 | 6  | AC149597 Mus muscu             |
| 23 | 77.4 | 4.1 | 153266 | 6  | AC165336 Mus muscu             |
| 24 | 76.2 | 4.1 | 129295 | 12 | AC178570 Strongylo             |
| 25 | 74.6 | 4.0 | 7218   | 2  | I66494 Sequence 14             |
| 26 | 74.6 | 4.0 | 16377  | 13 | AF002180 Plasmodu              |
| 27 | 74.4 | 4.0 | 233713 | 12 | AC133034 Rattus no             |
| 28 | 73.6 | 3.9 | 111554 | 4  | AP002460 Arabidops             |
| 29 | 73   | 3.9 | 271546 | 13 | AE014843 Plasmodu              |
| 30 | 72.2 | 3.8 | 5420   | 13 | AF056936 Plasmodu              |
| 31 | 71   | 3.8 | 110000 | 15 | CP000061_3 Continuation (5 of  |
| 32 | 71   | 3.8 | 110000 | 15 | CP000061_4 Continuation (5 of  |
| 33 | 71   | 3.8 | 175544 | 12 | AC117342 Rattus no             |
| 34 | 70.6 | 3.8 | 66442  | 13 | PFMAL1P4 Arabidops             |
| 35 | 70.4 | 3.7 | 1686   | 2  | E08995 DNA encodin             |
| 36 | 70   | 3.7 | 110000 | 15 | CP000111_04 Continuation (5 of |
| 37 | 70   | 3.7 | 257757 | 13 | AE014837 Plasmodu              |
| 38 | 69.8 | 3.7 | 110000 | 13 | PFMAL1P2_0 Arabidops           |
| 39 | 69.8 | 3.7 | 348174 | 13 | CR382399 Plasmodu              |
| 40 | 69.4 | 3.7 | 208332 | 6  | AC122227 Mus muscu             |
| 41 | 69.4 | 3.7 | 206187 | 12 | AC182450 Mus muscu             |
| 42 | 69.2 | 3.7 | 825    | 13 | AF206632 Plasmodu              |
| 43 | 69.2 | 3.7 | 81019  | 4  | AB073159 Arabidops             |
| 44 | 69.2 | 3.7 | 253924 | 13 | AE014822 Plasmodu              |
| 45 | 68.6 | 3.6 | 4995   | 13 | PFMMSA M69183 Plasmodu         |

#### ALIGNMENTS

RESULT 1  
AR396704 1980 bp DNA linear PAT 18-DEC-2003  
LOCUS Sequence 2719 from patent US 6617156.  
DEFINITION AR396704  
ACCESSION AR396704  
VERSION AR396704.1 GI:40126444  
KEYWORDS  
SOURCE  
ORGANISM  
Unknown.  
REFERENCE  
1 (bases 1 to 1980)  
Doucette-Stamm, L.A. and Bush, D.  
Nucleic acid and amino acid sequences relating to Enterococcus  
faecalis for diagnostics and therapeutics  
JOURNAL Patent: US 6617156-A 2719 09-SBP-2003;  
FEATURES  
location/Qualifiers  
1..1980  
/organism="unknown"  
/mol\_type="genomic DNA"

#### ORIGIN

Query Match 100.0%; Score 1881; DB 2; Length 1980;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1881; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ATGAAGCAATTAATAAAAGTTGGTACACCGTTAGTCTGTACTAATTTGCCACTT 60  
DB 97 ATGAAGCAATTAATAAAAGTTGGTACACCGTTAGTCTGTACTAATTTGCCACTT 156  
QY 61 TTCACAAGCTTTTGGGACACCACTGATTTGAGAAAGAAATGGGAGCGCACAG 120  
DB 157 TTCACAAGCTTTTGGGACACCACTGATTTGAGAAAGAAATGGGAGCGCACAG 216  
QY 121 CTCGATTTCAAAAGAAATGACGATTTACAGATCCGCTTATTCAAATATGCGGG 180  
DB 217 CTCGATTTCAAAAGAAATGACGATTTACAGATCCGCTTATTCAAATATGCGGG 276  
QY 181 AAAAGAAATGACGATTTGATTAATATCAAGACATGAGATGATGACGTTTGTATTTAT 240  
DB 277 AAAAGAAATGACGATTTGATTAATATCAAGACATGAGATGATGACGTTTGTATTTAT 336

241 AACGTGACGAAGATTTTACGAGCAACGAGCGGCGCAAGCGTTGTCAGCTAA 300  
| | | | |  
Db 337 AACGTGACGAAGATTTTACGAGCAACGAGCGGCGCAAGCGTTGTCAGCTAA 396  
| | | | |  
Qy 301 CAAGCTGTCGAAGATTAACTCTGGGAAAACCTGTTGCTCAAGAAACACCGATGCAAT 360  
| | | | |  
Db 397 CAAGCTGTCGAAGATTAACTCTGGGAAAACCTGTTGCTCAAGAAACACCGATGCAAT 456  
| | | | |  
Qy 361 GGAATGTCAGTTCAGTTACCTTAAATAAATAAGTAAAGATGCAAGTATACCAT 420  
| | | | |  
Db 457 GGAATGTCAGTTCAGTTACCTTAAATAAATAAGTAAAGATGCAAGTATACCAT 516  
| | | | |  
Qy 421 AAAGAAGAACCAAAAGAGGTGATGCTGCTACCAATATGAGTGGCGTTCCAGTT 480  
| | | | |  
Db 517 AAAGAAGAACCAAAAGAGGTGATGCTGCTACCAATATGAGTGGCGTTCCAGTT 576  
| | | | |  
Qy 481 TACGAATGATCAAGCAAAACAGATGTTCTTATTAATATGAAACAGAGATTAAGCGTT 540  
| | | | |  
Db 577 TACGAATGATCAAGCAAAACAGATGTTCTTATTAATATGAAACAGAGATTAAGCGTT 636  
| | | | |  
Qy 541 GTTCATATTTATCTTAAATATGATGAGCAATGATGATGTTACATGTAATAAAGTA 600  
| | | | |  
Db 637 GTTCATATTTATCTTAAATATGATGAGCAATGATGATGTTACATGTAATAAAGTA 696  
| | | | |  
Qy 601 GGAAGCTGCTGAATATGAAGATTAATATGAGCAATGTTGTTATTTCTAAAGCGAAGC 660  
| | | | |  
Db 697 GGAAGCTGCTGAATATGAAGATTAATATGAGCAATGTTGTTATTTCTAAAGCGAAGC 756  
| | | | |  
Qy 661 TCACCAAGGCAAGTAAATATATCCAGAGTCAAAAGATGATTAATATCATGAGCAAG 720  
| | | | |  
Db 757 TCACCAAGGCAAGTAAATATATCCAGAGTCAAAAGATGATTAATATCATGAGCAAG 816  
| | | | |  
Qy 721 GATTAAGAACCAAGCAAAAGCTTTATTAATGAGGAAAATTATGAATTTGGCGAAATGAT 780  
| | | | |  
Db 817 GATTAAGAACCAAGCAAAAGCTTTATTAATGAGGAAAATTATGAATTTGGCGAAATGAT 876  
| | | | |  
Qy 781 TTCAAGAGAGCAGAGATGAAACGGGAGATTAACAGTTAAATCTTGAGGTGGTTCG 840  
| | | | |  
Db 877 TTCAAGAGAGCAGAGATGAAACGGGAGATTAACAGTTAAATCTTGAGGTGGTTCG 936  
| | | | |  
Qy 841 TATATTTTGAAGAAAGTAAAGCTCCAAATATGCAATTAATGAAATCAAACAAA 900  
| | | | |  
Db 937 TATATTTTGAAGAAAGTAAAGCTCCAAATATGCAATTAATGAAATCAAACAAA 996  
| | | | |  
Qy 901 ACACCATTTTACATTTGAGCAAAACATCAACACCTGTTGAAAACCAAGTCAAAATGAT 960  
| | | | |  
Db 997 ACACCATTTTACATTTGAGCAAAACATCAACACCTGTTGAAAACCAAGTCAAAATGAT 1056  
| | | | |  
Qy 961 ACCTCTAAAGTTGATTAACCAACACCAAGCTTAGATGTTAAAGATGTGCAATTGGCGAA 1020  
| | | | |  
Db 1057 ACCTCTAAAGTTGATTAACCAACACCAAGCTTAGATGTTAAAGATGTGCAATTGGCGAA 1116  
| | | | |  
Qy 1021 AAAATTAATTAATCAAAATTTCTGTAATATTTCCATTTGGGATTTGCAGACAAAAGGCGAC 1080  
| | | | |  
Db 1117 AAAATTAATTAATCAAAATTTCTGTAATATTTCCATTTGGGATTTGCAGACAAAAGGCGAC 1176  
| | | | |  
Qy 1081 GCTAATTAATATCGTCAATTTCAATTTAGTTGATTAACATGATGAGCTTTAATCTTTGAT 1140  
| | | | |  
Db 1177 GCTAATTAATATCGTCAATTTCAATTTAGTTGATTAACATGATGAGCTTTAATCTTTGAT 1236  
| | | | |  
Qy 1141 AACGTGACTTTCGAGAGATATGCTTATATGATGAGGAGATACAGTGAATGCTCT 1200  
| | | | |  
Db 1237 AACGTGACTTTCGAGAGATATGCTTATATGATGAGGAGATACAGTGAATGCTCT 1296  
| | | | |  
Qy 1201 GAAATTTTCAAGTATCAAGCAAAATGCTTCACTGTCGCGCTTAATCAAGCTAT 1260  
| | | | |  
Db 1297 GAAATTTTCAAGTATCAAGCAAAATGCTTCACTGTCGCGCTTAATCAAGCTAT 1356  
| | | | |  
Qy 1261 ATTCTTACGCTTAAACCAAGCGGCAACATTAATTTGTTACTTTATGCAATTTAATGAA 1320  
| | | | |  
Db 1357 ATTCTTACGCTTAAACCAAGCGGCAACATTAATTTGTTACTTTATGCAATTTAATGAA 1416  
| | | | |  
Qy 1321 AAAGCAGATCTTACGAAGAGCTTTAAATAATGAGGCGAATGTTGATTAACGCTATACCGAC 1380  
| | | | |

1417 AAAGCAGATCTTACGAAGAGCTTTAAATAATGAGGGAATGTTGATTAAGCTATACCGAC 1476  
| | | | |  
Db 1381 GACCAAAACCAACCAACCTTTGAAGTTTGAACAGTGGGAAACCTTTCACTTAAAGTCAT 1440  
| | | | |  
Qy 1477 GACCAAAACCAACCAACCTTTGAAGTTTGAACAGTGGGAAACCTTTCACTTAAAGTCAT 1536  
| | | | |  
Db 1441 GGCAGATGTCAGAGCGCAACAAGCCTTTGGGAGGCTTCTTTGTCGTCGATCAAAAC 1500  
| | | | |  
Qy 1537 GGCAGATGTCAGAGCGCAACAAGCCTTTGGGAGGCTTCTTTGTCGTCGATCAAAAC 1596  
| | | | |  
Db 1501 AGCGACACGCAAAATTTATTTGAAATTCATGAAACACGAAAGCAGCAACTTGGTGAAA 1560  
| | | | |  
Qy 1597 AGCGACACGCAAAATTTATTTGAAATTCATGAAACACGAAAGCAGCAACTTGGTGAAA 1656  
| | | | |  
Db 1561 ACMAAAGCTGAAGCAACCTTCTTCAACAAGGCTGATGATTAATGATTCACAGG 1620  
| | | | |  
Qy 1657 ACMAAAGCTGAAGCAACCTTCTTCAACAAGGCTGATGATTAATGATTCACAGG 1716  
| | | | |  
Db 1621 CTTAAATACGCTACCTATTTATTAAGAAACCTGATGCTCTGATGATTAATGCTTTGTA 1680  
| | | | |  
Qy 1717 CTTAAATACGCTACCTATTTATTAAGAAACCTGATGCTCTGATGATTAATGCTTTGTA 1776  
| | | | |  
Db 1681 ACMAATCGATTTGAATTTGCTGCTCAATGAAACATCATATGSCAACACAGAAAACCTGATT 1740  
| | | | |  
Qy 1777 ACMAATCGATTTGAATTTGCTGCTCAATGAAACATCATATGSCAACACAGAAAACCTGATT 1836  
| | | | |  
Db 1741 TCACCAAGAAAGTACCAACCAACCAAGGCTTACCTTCAACAGGTGGCAAGAA 1800  
| | | | |  
Qy 1837 TCACCAAGAAAGTACCAACCAACCAAGGCTTACCTTCAACAGGTGGCAAGAA 1896  
| | | | |  
Db 1801 ATCTACGTTTACTTGAAGTGGCGCAGCTTGTCTACTTATTCAGAGTCTTACTTGTCT 1860  
| | | | |  
Qy 1897 ATCTACGTTTACTTGAAGTGGCGCAGCTTGTCTACTTATTCAGAGTCTTACTTGTCT 1956  
| | | | |  
Db 1861 AGACGTAGAAAAGAAATGCT 1881  
| | | | |  
Qy 1957 AGACGTAGAAAAGAAATGCT 1977  
| | | | |  
Db  
  
RESULT 2  
BD191807  
LOCUS BD191807 1983 bp DNA linear PART 17-JUL-2003  
DEFINITION Enterococcus Faecalis polynucleotides and polypeptides.  
ACCESSION BD191807  
VERSION BD191807.1 GI:33001546  
KEYWORDS JP 2002516571-A/109.  
SOURCE synthetic construct  
ORGANISM synthetic construct  
REFERENCE 1 (bases 1 to 1983)  
AUTHORS Kunisch, C.A., Choi, G.H., Bailey, C. and Hromockyj, A.  
TITLE Enterococcus Faecalis polynucleotides and polypeptides  
JOURNAL Patent: JP 2002516571-A 109 04-JUN-2002;  
HUMAN GENOME SCIENCES INC  
COMMENT  
PN JP 2002516571-A/109  
PD 04-JUN-2002  
PF 04-MAY-1998 JP 1998548289  
PR 06-MAY-1997 US 60/044031, 16-MAY-1997 US 60/046655 PR  
14-NOV-1997 US 60/066009  
PI CHARLES A KUNSCH, GIL H CHOI, CAMELLA BAILEY, ALEX HROMOCKYJ PC  
C12N15/31, C07K14/315, C07K16/12, C12Q1/68, C12N1/21, C12N5/12, PC  
G01N33/569,  
PC G01N33/68, A61K39/09  
CC Strandedness: Double;  
CC Topology: Linear;  
FH Key Location/Qualifiers.  
FEATURES  
source location/Qualifiers  
1..1983  
/organism="synthetic construct"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:32630"  
ORIGIN

Query Match 100.0%; Score 1881; DB 2; Length 1983;  
 Beet Local Similarity 100.0%; Pred. No. 0;  
 Matches 1881; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 ATGAAGCAATTAATAAAGTTTGGTACACCGTATGATCTGTATTAATTTTGGCACTT 60  
 100 ATGAAGCAATTAATAAAGTTTGGTACACCGTATGATCTGTATTAATTTTGGCACTT 159  
 61 TTCAACAAGTGTATTTAGGACAAACAATGCAATTTGGAGAAAGAGAGAGAGAG 120  
 160 TTCAACAAGTGTATTTAGGACAAACAATGCAATTTGGAGAAAGAGAGAGAGAG 219  
 121 CTCGATTTCACAAAAAGAAATGAGAGATTTACCAAGTCCGCTTATTCAAAATAGCGG 180  
 220 CTCGATTTCACAAAAAGAAATGAGAGATTTACCAAGTCCGCTTATTCAAAATAGCGG 279  
 181 AAGAAGATGAGAGATTTGATTAATATCAAGGACTGGCAGATGTGACGTTTATTTAT 240  
 280 AAGAAGATGAGAGATTTGATTAATATCAAGGACTGGCAGATGTGACGTTTATTTAT 339  
 241 AAGTGAAGCAAGATTTTACAGACAAAGAGCGGCGCAAGCGTTGATGCACTTAA 300  
 340 AAGTGAAGCAAGATTTTACAGACAAAGAGCGGCGCAAGCGTTGATGCACTTAA 399  
 301 CAAGCTGTCCAAAGTTTAACTCGTGGGAAACCTGTGCTCAAGAAACCAAGATGCAAT 360  
 400 CAAGCTGTCCAAAGTTTAACTCGTGGGAAACCTGTGCTCAAGAAACCAAGATGCAAT 459  
 361 GGAAGTGTCACTGTTCAGTTTACCTTAATAAACAATGATAAGATGCACTTATACAT 420  
 460 GGAAGTGTCACTGTTCAGTTTACCTTAATAAACAATGATAAGATGCACTTATACAT 519  
 421 AAGAAGAAACCAAAAGAGGCTGTAGTGTCTACGAATATGTGTGCGTTCCAGTT 480  
 520 AAGAAGAAACCAAAAGAGGCTGTAGTGTCTACGAATATGTGTGCGTTCCAGTT 579  
 481 TAAGAAATGATCAAGAAACAGATGCTCTTAATAATATGAAACGAAGAAATTAAGCGGT 540  
 580 TAAGAAATGATCAAGAAACAGATGCTCTTAATAATATGAAACGAAGAAATTAAGCGGT 639  
 541 GTTCATATTTATCTTAATAATGATGTAGCCATGATGATGTTTACATGTGAAAAAGTA 600  
 640 GTTCATATTTATCTTAATAATGATGTAGCCATGATGATGTTTACATGTGAAAAAGTA 699  
 601 GGAATCTGTGAATAAGATTTAAATGCGCAGAAATTTGTTTCTTAAAGCGAAGGC 660  
 700 GGAATCTGTGAATAAGATTTAAATGCGCAGAAATTTGTTTCTTAAAGCGAAGGC 759  
 661 TCACAGAGCAAGTAAATATATCCAGAGATCAAGATGATTTATATCATGACAAAG 720  
 760 TCACAGAGCAAGTAAATATATCCAGAGATCAAGATGATTTATATCATGACAAAG 819  
 721 GATAAAGAAACAAGAAACGCTTTATTAATCTGGGAAAAAGTTAGAAATTTGGCAAAATGAT 780  
 820 GATAAAGAAACAAGAAACGCTTTATTAATCTGGGAAAAAGTTAGAAATTTGGCAAAATGAT 879  
 781 TTCAACAAGCAAGAAATGAAACGCGAGAAATTAACAGTTTAAATCTTGAAGTTGCTTCG 840  
 880 TTCAACAAGCAAGAAATGAAACGCGAGAAATTAACAGTTTAAATCTTGAAGTTGCTTCG 939  
 841 TATATTTTGAAGAAAGTAAAGCTCAATTAATGCAAGATTTAATGAAATCAACAA 900  
 940 TATATTTTGAAGAAAGTAAAGCTCAATTAATGCAAGATTTAATGAAATCAACAA 999  
 901 ACACCAATTTACATTTGAAGCAACATCAACACCTGTGTAAGAAAAACAGTCAAAATGAT 960  
 1000 ACACCAATTTACATTTGAAGCAACATCAACACCTGTGTAAGAAAAACAGTCAAAATGAT 1059  
 961 ACCTTAAGTTGATTAACAAACACCAAGCTTATGATGTAAGATGTGCAATTTGGCGAA 1020  
 1060 ACCTTAAGTTGATTAACAAACACCAAGCTTATGATGTAAGATGTGCAATTTGGCGAA 1119

QY 1021 AAAATTTAAATCAATTTCTGTAAATATTCATTTGGGGATTCAGACAAAGAGCGAC 1080  
 DB 1120 AAAATTTAAATCAATTTCTGTAAATATTCATTTGGGGATTCAGACAAAGAGCGAC 1179  
 QY 1081 GCTAAATTAATACGTCAATTTCAATTTAGTTGATTAACAATGATGACGCTTAACTTTTGTAT 1140  
 DB 1180 GCTAAATTAATACGTCAATTTCAATTTAGTTGATTAACAATGATGACGCTTAACTTTTGTAT 1239  
 QY 1141 AAGCTGACTTCTGAGAGATTAATGCTTATATGATGGGATTAAGTATGCTCTCT 1200  
 DB 1240 AAGCTGACTTCTGAGAGATTAATGCTTATATGATGGGATTAAGTATGCTCTCTCT 1299  
 QY 1201 GAAAATTAATCAAGTACGTAACAAGCAATGAGCTTCACTGTGCGCTTAAATCAAGCTAT 1260  
 DB 1300 GAAAATTAATCAAGTACGTAACAAGCAATGAGCTTCACTGTGCGCTTAAATCAAGCTAT 1359  
 QY 1261 ATTCTACGCTTAACACAGCGCGCACTAAATTCGTTTACTTTATGCAATTTAAATGAA 1320  
 DB 1360 ATTCTACGCTTAACACAGCGCGCACTAAATTCGTTTACTTTATGCAATTTAAATGAA 1419  
 QY 1321 AAGCAATCTTAACGAAGGCTTTAAATAAGCGAATGTTGATTAACGCTATACGAC 1380  
 DB 1420 AAGCAATCTTAACGAAGGCTTTAAATAAGCGAATGTTGATTAACGCTATACGAC 1479  
 QY 1381 GACCAAAACACCAACTGTTGAAGTTGACAGGTTGGGAAACGTTTCAATTAAGTCGAT 1440  
 DB 1480 GACCAAAACACCAACTGTTGAAGTTGACAGGTTGGGAAACGTTTCAATTAAGTCGAT 1539  
 QY 1441 GCGATGTGACAGCAACAAACCTTGGCGGAGCTTCTTGTGCTCCGATCAAAAC 1500  
 DB 1540 GCGATGTGACAGCAACAAACCTTGGCGGAGCTTCTTGTGCTCCGATCAAAAC 1599  
 QY 1501 AGCGACACGAATTTATTTGAAATTCATGAAACAGAAACGCAACCTTGGGTGAAA 1560  
 DB 1600 AGCGACACGAATTTATTTGAAATTCATGAAACAGAAACGCAACCTTGGGTGAAA 1659  
 QY 1561 ACAAAAGCTGAAGCAACTTCTTCAACAAAGCGTATGATTTGATTAATCAAGG 1620  
 DB 1660 ACAAAAGCTGAAGCAACTTCTTCAACAAAGCGTATGATTTGATTAATCAAGG 1719  
 QY 1621 CTTAAATACGTAACCTTATTTAGAGAAACCTGAGCTCTGATGATTTATCTTGTGA 1680  
 DB 1720 CTTAAATACGTAACCTTATTTAGAGAAACCTGAGCTCTGATGATTTATCTTGTGA 1779  
 QY 1681 ACAATTCGATTAATTTGTGCTCAATGACATATATGCGCAACAGAAACCTAGTT 1740  
 DB 1780 ACAATTCGATTAATTTGTGCTCAATGACATATATGCGCAACAGAAACCTAGTT 1839  
 QY 1741 TCACAGAAAAAGTACCAACAAACAAAGGATCTTAACCTTCAACAGGTGCAAAAGGA 1800  
 DB 1840 TCACAGAAAAAGTACCAACAAACAAAGGATCTTAACCTTCAACAGGTGCAAAAGGA 1899  
 QY 1801 ATCTAGTTTACTTGAAGTGGCGCAGTCTTGTACTTATTTGAGAGATCTTACTTGT 1860  
 DB 1900 ATCTAGTTTACTTGAAGTGGCGCAGTCTTGTACTTATTTGAGAGATCTTACTTGT 1959  
 QY 1861 AGAGCTAGAAAAAAGAAATGCT 1881  
 DB 1960 AGAGCTAGAAAAAAGAAATGCT 1980

RESULT 3  
 LOCUS AR228552 1983 bp DNA linear PAT 20-DEC-2002  
 DEFINITION Sequence 217 from patent US 6448043.  
 ACCESSION AR228552  
 VERSION AR228552.1 GI:27267389  
 KEYWORDS  
 SOURCE Unknown.  
 ORGANISM Unknown.  
 REFERENCE 1 (bases 1 to 1983)  
 Chou, G.H., Bailey, C., Hromockyj, A. and Kunsch, C.A.

TITLE Enterococcus faecalis EF040 and uses therefor  
JOURNAL Patent: US 6448043-A 217 10-SEP-2002;  
Human Genome Sciences, Inc.; Rockville, MD;

FEATURES  
source Location/Qualifiers

1..1983  
/organism="unknown"  
/mol\_type="genomic DNA"

## ORIGIN

Query Match 100.0%; Score 1881; DB 2; Length 1983;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1881; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 ATGAAGCAATTAAAAAGTTGGTACACCGTTAGTACCTTGTACTAATTTTGCACATT 60
DB 100 ATGAAGCAATTAAAAAGTTGGTACACCGTTAGTACCTTGTACTAATTTTGCACATT 159
QY 61 TTCACAAAGTGTATTAGGACAACAACACTGCTTGCAGAGAAAATGGGGAGACGACAG 120
DB 160 TTCACAAAGTGTATTAGGACAACAACACTGCTTGCAGAGAAAATGGGGAGACGACAG 219
QY 121 CTCGTGATTACAAAAAGAAAATGACGGAATTTACAGATCCGCTTATTCAAAATAGCGG 180
DB 220 CTCGTGATTACAAAAAGAAAATGACGGAATTTACAGATCCGCTTATTCAAAATAGCGG 279
QY 181 AAGAGAAATAGCGAGCTTTGATTAATATCAAGGACTGGGAGATGTGACGTTTATGATTTAT 240
DB 280 AAGAGAAATAGCGAGCTTTGATTAATATCAAGGACTGGGAGATGTGACGTTTATGATTTAT 339
QY 241 AACGTGACAAAGCAATTTTACGACGACGCGGCAAGCGCAAGCGTTGATGCACTAAA 300
DB 340 AACGTGACAAAGCAATTTTACGACGACGCGGCAAGCGCGCAAGCGTTGATGCACTAAA 399
QY 301 CAAGCTGTCCAAAGTTTAACTCTCGGGAAAACCTGTTGCTCAAGGAAACCAACGATGCAAT 360
DB 400 CAAGCTGTCCAAAGTTTAACTCTCGGGAAAACCTGTTGCTCAAGGAAACCAACGATGCAAT 459
QY 361 GGGAAATGCACTGTTCAGTTTAACTTAAAAAACAAGGTAAAGTCAAGATGATTAACATT 420
DB 460 GGGAAATGCACTGTTCAGTTTAACTTAAAAAACAAGGTAAAGTCAAGATGATTAACATT 519
QY 421 AAGAGAAACCAAAAGAGGGTGTAGTGTCTACGAAATGTGTGTGCGTTCCAGTT 480
DB 520 AAGAGAAACCAAAAGAGGGTGTAGTGTCTACGAAATGTGTGTGCGTTCCAGTT 579
QY 481 TAGGAATGATCAAGCAACAGATGCTTCTATTAATATGAAACAGAGAAATTAAGCGTT 540
DB 580 TAGGAATGATCAAGCAACAGATGCTTCTATTAATATGAAACAGAGAAATTAAGCGTT 639
QY 541 GTTCATATTATTCCTTAAAAATGTGTAGGCAATGATGTATGTTAATGAAAAAGTA 600
DB 640 GTTCATATTATTCCTTAAAAATGTGTAGGCAATGATGTATGTTAATGAAAAAGTA 699
QY 601 GGAAGTGTGAAAAATGAGGATTTAAATGGCGCAGAAATTTGTTATTTCTAAAAAGCGAAGGC 660
DB 700 GGAAGTGTGAAAAATGAGGATTTAAATGGCGCAGAAATTTGTTATTTCTAAAAAGCGAAGGC 759
QY 661 TCACCAAGGCAAGTAAATATATTCAGAGATCAAGAGTCAAGATGATTAATATCATGACACAG 720
DB 760 TCACCAAGGCAAGTAAATATATTCAGAGATCAAGAGTCAAGATGATTAATATCATGACACAG 819
QY 721 GATTAAGAAACAAAGCAAAAGCTTTATTAATCTGGGAAAAGTTATGAAATTTGGCGAAATGAT 780
DB 820 GATTAAGAAACAAAGCAAAAGCTTTATTAATCTGGGAAAAGTTATGAAATTTGGCGAAATGAT 879
QY 781 TTCACAGAAAGCAGAGATGAAACGGGAGATTTAAACGTTAAAAATCTTGAGTTGGTTG 840
DB 880 TTCACAGAAAGCAGAGATGAAACGGGAGATTTAAACGTTAAAAATCTTGAGTTGGTTG 939
QY 841 TATATTTTGAAGAGTAAAGCTCCAAATATGCAATTAATGAAATTAACAAACAAA 900
DB 940 TATATTTTGAAGAGTAAAGCTCCAAATATGCAATTAATGAAATTAACAAACAAA 999
```

```
QY 901 AACACATTTTACAATTGAAGCAAAACAATCAAAACCTGTTGAAAAAAGTCAAAAATGAT 960
DB 1000 AACACATTTTACAATTGAAGCAAAACAATCAAAACCTGTTGAAAAAAGTCAAAAATGAT 1059
QY 961 ACCCTTAAAGTTGATTAACCAACACAGCTTGAATGTAAGATGTGCAATTTGGCGAA 1020
DB 1060 ACCCTTAAAGTTGATTAACCAACACAGCTTGAATGTAAGATGTGCAATTTGGCGAA 1119
QY 1021 AAAATTTAATATCAAAATTTCTGTAAATATTCATTTGGGGATTCAGACAAAGAGCGAC 1080
DB 1120 AAAATTTAATATCAAAATTTCTGTAAATATTCATTTGGGGATTCAGACAAAGAGCGAC 1179
QY 1081 GCTAATTAATAGTCAAAATTCATTTAGTTGATTAACATGATGACGCTTAACTTTGAT 1140
DB 1180 GCTAATTAATAGTCAAAATTCATTTAGTTGATTAACATGATGACGCTTAACTTTGAT 1239
QY 1141 AACGTGACTTCTGGAGAGATGCTTATATGCTTATATGATGAGGAGATACAGTGAATGCTCCT 1200
DB 1240 AACGTGACTTCTGGAGAGATGCTTATATGCTTATATGATGAGGAGATACAGTGAATGCTCCT 1299
QY 1201 GAAAATTTATCAAGTGACTGAACACAGAAATGGCTTCACTGTGCGCTTATATCAGCGTAT 1260
DB 1300 GAAAATTTATCAAGTGACTGAACACAGCAATGGCTTCACTGTGCGCTTATATCAGCGTAT 1359
QY 1261 ATTCCAGGCTTAAACAAGGCGGCAACCTAAATTCGTTTACTTTATGCAATTTAAATGAA 1320
DB 1360 ATTCCAGGCTTAAACAAGGCGGCAACCTAAATTCGTTTACTTTATGCAATTTAAATGAA 1419
QY 1321 AAGAGCAATCTTACGAAAGGCTTTTAAAAATGAGGCGAATGTTATTAACGCTATACCGAC 1380
DB 1420 AAGAGCAATCTTACGAAAGGCTTTTAAAAATGAGGCGAATGTTATTAACGCTATACCGAC 1479
QY 1381 GACCAAAACCAACCAACTGTTGAATGTGACAGGTGGGAAAACGTTTCACTTAAATGCAAT 1440
DB 1480 GACCAAAACCAACCAACTGTTGAATGTGACAGGTGGGAAAACGTTTCACTTAAATGCAAT 1539
QY 1441 GGGAGATGACAGCGCACACAACTTGGCGGAGGCTTCTTTGCTGCTCGATCAAAAC 1500
DB 1540 GGGAGATGACAGCGCACACAACTTGGCGGAGGCTTCTTTGCTGCTCGATCAAAAC 1599
QY 1501 AGCGACACAGCAAAATTTATGAAAAATCGATGAAACAAACGAAACGACCAACTTGGGTGAAA 1560
DB 1600 AGCGACACAGCAAAATTTATGAAAAATCGATGAAACAAACGAAACGACCAACTTGGGTGAAA 1659
QY 1561 ACAAAGCTGGAAGCAACTTCTTTTACAACAAGCTGATGATGTTGATATCAACAGG 1620
DB 1660 ACAAAGCTGGAAGCAACTTCTTTTACAACAAGCTGATGATGTTGATATCAACAGG 1719
QY 1621 CTTAAATACGTAACCTATTAATTTAGAAAGAACTGTAGCTCTGATGATTAATGCTTTGTA 1680
DB 1720 CTTAAATACGTAACCTATTAATTTAGAAAGAACTGTAGCTCTGATGATTAATGCTTTGTA 1779
QY 1681 ACAAAATCGGATTTGAATTTGTGTCAATGAAACATATATGCAACACAGAAAACTTATGTT 1740
DB 1780 ACAAAATCGGATTTGAATTTGTGTCAATGAAACATATATGCAACACAGAAAACTTATGTT 1839
QY 1741 TCACCAAGAAAGTACCAAAACCAACCAAGGTAACCTTCAACAGATGCGCAAAAGGA 1800
DB 1840 TCACCAAGAAAGTACCAAAACCAACCAAGGTAACCTTCAACAGATGCGCAAAAGGA 1899
QY 1801 ATCTACGTTTACTTGAAGAGTGGCGCAGCTTGTCTACTTATTTGACAGAGTCTACTTTGCT 1860
DB 1900 ATCTACGTTTACTTGAAGAGTGGCGCAGCTTGTCTACTTATTTGACAGAGTCTACTTTGCT 1959
QY 1861 AGACGTAGAAAAAATAATGCT 1881
DB 1960 AGACGTAGAAAAAATAATGCT 1980
```

RESULT 4  
AR692967  
LOCUS AR692967 1983 bp DNA linear PAT 14-SEP-2005



DEFINITION Sequence 217 from patent US 6913907.  
ACCESSION AR692967  
VERSION AR692967.1 GI:75182710  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 1983)  
AUTHORS Choi, G.H., Bailey, C., Hromocky, A. and Kunsch, C.A.  
TITLE Enterococcus faecalis polynucleotides encoding EF059  
JOURNAL Patent: US 6913907-A 217 05-JUL-2005;  
Human Genome Sciences, Inc.; Rockville, MD;  
WOM;

FEATURES  
source Location/Qualifiers  
1..1983  
/organism="unknown"  
/mol\_type="genomic DNA"

ORIGIN

Query Match 100.0%; Score 1881; DB 2; Length 1983;  
Best Local Similarity 100.0%; Pred.No. 0;  
Matches 1881; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAAGCAATTAATAAAGTTGGTACACCGTAGTACCTGTTACTAATTTTGGCAGCTT 60  
DB 100 ATGAAGCAATTAATAAAGTTGGTACACCGTAGTACCTGTTACTAATTTTGGCAGCTT 159  
QY 61 TTCACAGTGTATTAGGGAACAACAACGCTGATTTGAGAAAGAAATGGGAGAGCGCAG 120  
DB 160 TTCACAGTGTATTAGGGAACAACAACGCTGATTTGAGAAAGAAATGGGAGAGCGCAG 219  
QY 121 CTCGTGATTCACAAAAGAAATGACGAGATTACCGATCCGTTATTCAAATAGCGGG 180  
DB 220 CTCGTGATTCACAAAAGAAATGACGAGATTACCGATCCGTTATTCAAATAGCGGG 279  
QY 181 AAAGAAATGAGCGAGTTGTATAATATCAAGGACTGGCAGATGTGACGTTTGTATTTAT 240  
DB 280 AAAGAAATGAGCGAGTTGTATAATATCAAGGACTGGCAGATGTGACGTTTGTATTTAT 339  
QY 241 AACGTGACGAACGAATTTTACGACCAACGAGCGGCGCAAGCGTTGATGACGCTAA 300  
DB 340 AACGTGACGAACGAATTTTACGACCAACGAGCGGCGCAAGCGTTGATGACGCTAA 399  
QY 301 CAAGCTGTCCAAAGTTTAACTCTCGGGAACCTGTGCTCAAGAAACCAAGTGCAT 360  
DB 400 CAAGCTGTCCAAAGTTTAACTCTCGGGAACCTGTGCTCAAGAAACCAAGTGCAT 459  
QY 361 GGGAAATGTCAGTGTCACTTAACAAAACAAATGTAAGATGATGATATACATT 420  
DB 460 GGGAAATGTCAGTGTCACTTAACAAAACAAATGTAAGATGATGATATACATT 519  
QY 421 AAAGAAACCAAAAGAGGGGTAGTGTGCTACGAATATGATGATGATGATGATGAT 480  
DB 520 AAAGAAACCAAAAGAGGGGTAGTGTGCTACGAATATGATGATGATGATGATGAT 579  
QY 481 TACGAAATGATCAAGCAACAGATGTTTCTTAATTAATGAAACAGAAATTAAGGGT 540  
DB 580 TACGAAATGATCAAGCAACAGATGTTTCTTAATTAATGAAACAGAAATTAAGGGT 639  
QY 541 GTTCAATATTTATCTTAATAATGTGTAGCAATGATGATGATGATGATGATGATGAT 600  
DB 640 GTTCAATATTTATCTTAATAATGTGTAGCAATGATGATGATGATGATGATGATGAT 699  
QY 601 GGAATGCTGAATAATGAGGATTAATGCGCAAGATTTGTTATTTCTAAAGCGAAGGC 660  
DB 700 GGAATGCTGAATAATGAGGATTAATGCGCAAGATTTGTTATTTCTAAAGCGAAGGC 759  
QY 661 TCACACGAGCAAGTAAATATATCAAGAGTCAAGATGATGATGATGATGATGATGATGAT 720  
DB 760 TCACACGAGCAAGTAAATATATCAAGAGTCAAGATGATGATGATGATGATGATGATGAT 819  
QY 721 GATTAAGAACAAAGCAAAACGCTTATTAATGAGAAAGTTAATGAAATGGCAAAATGAT 780

DB 820 GATTAAGAACAAAGCAAAACGCTTATTAATGAGAAAGTTAATGAAATGGCAAAATGAT 879  
QY 781 TTCACAGAAACAGAGATGAAACGGGAAATTAACGTTAAATCTTGAAGTTGCTCG 840  
DB 880 TTCACAGAAACAGAGATGAAACGGGAAATTAACGTTAAATCTTGAAGTTGCTCG 939  
QY 841 TATATTTTGAAGAAAGTAAAGCTCCAAATTAATGCAAGTAAATGAAATCAAAACAAA 900  
DB 940 TATATTTTGAAGAAAGTAAAGCTCCAAATTAATGCAAGTAAATGAAATCAAAACAAA 999  
QY 901 ACACCATTTACATTTAGAACCAACATCAACACCTGTGAAAGAAACAGTCAAAATGAT 960  
DB 1000 ACACCATTTACATTTAGAACCAACATCAACACCTGTGAAAGAAACAGTCAAAATGAT 1059  
QY 961 ACCTCTAAAGTTGATTAACCAACACCAAGCTTATGATGATGATGATGATGATGATGAT 1020  
DB 1060 ACCTCTAAAGTTGATTAACCAACACCAAGCTTATGATGATGATGATGATGATGATGAT 1119  
QY 1021 AAAATTTAATTAATCAATTTCTGTTAAATATTCATTTGGGGAATGGCAACAAAGGCGAC 1080  
DB 1120 AAAATTTAATTAATCAATTTCTGTTAAATATTCATTTGGGGAATGGCAACAAAGGCGAC 1179  
QY 1081 GCTAATTAATTAATCAATTTCAATTTAGTTGATTAACATGATGACGCTTAACTTTGAT 1140  
DB 1180 GCTAATTAATTAATCAATTTCAATTTAGTTGATTAACATGATGACGCTTAACTTTGAT 1239  
QY 1141 AACGTGACTTCTGAGAGATATGCTTATATGATGATGATGATGATGATGATGATGATGAT 1200  
DB 1240 AACGTGACTTCTGAGAGATATGCTTATATGATGATGATGATGATGATGATGATGATGAT 1299  
QY 1201 GAAATTTAATCAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1260  
DB 1300 GAAATTTAATCAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1359  
QY 1261 ATTCTACGCTTAACACAGGCGGCAACATAAATTCGTTTACTTTATGATTAATGAA 1320  
DB 1360 ATTCTACGCTTAACACAGGCGGCAACATAAATTCGTTTACTTTATGATTAATGAA 1419  
QY 1321 AAAGCAATCTTAACGAAGGCTTTAAATTAATGAGCGAATGTTGATTAACGTTACGAC 1380  
DB 1420 AAAGCAATCTTAACGAAGGCTTTAAATTAATGAGCGAATGTTGATTAACGTTACGAC 1479  
QY 1381 GACCAAAACCAACCAACGTTGAAGTTGACAGGCGGGAACGTTTCAATTAAGTCAT 1440  
DB 1480 GACCAAAACCAACCAACGTTGAAGTTGACAGGCGGGAACGTTTCAATTAAGTCAT 1539  
QY 1441 GCGGATGTGACAGCAACAACGCTTGGCGGAGCTTCTTGTGCTCCGTGATCAAAAC 1500  
DB 1540 GCGGATGTGACAGCAACAACGCTTGGCGGAGCTTCTTGTGCTCCGTGATCAAAAC 1599  
QY 1501 AGCGACACGCAATTTATTTGAAATTCGATGAAACAGAAACGAAACCTTGGTGA 1560  
DB 1600 AGCGACACGCAATTTATTTGAAATTCGATGAAACAGAAACGAAACCTTGGTGA 1659  
QY 1561 ACAAAAGCTGAAGCAACTTCTTCAACAAAGGCGTATGATGATGATGATGATGATGATGAT 1620  
DB 1660 ACAAAAGCTGAAGCAACTTCTTCAACAAAGGCGTATGATGATGATGATGATGATGATGAT 1719  
QY 1621 CTTAAATACGTAACCTAATTTATTAAGAAACCTGATGCTCTGATGATTAATGCTTGT 1680  
DB 1720 CTTAAATACGTAACCTAATTTATTAAGAAACCTGATGCTCTGATGATTAATGCTTGT 1779  
QY 1681 ACAATTCGATTTGAATTTTGTGCTAATGAACATATGCAACAGAAACCTTAAT 1740  
DB 1780 ACAATTCGATTTGAATTTTGTGCTAATGAACATATGCAACAGAAACCTTAAT 1839  
QY 1741 TCACAGAAAAGTACCAACCAACAAAGGTAACCTTAACCTTCAACAGTGGCAAGGA 1800  
DB 1840 TCACAGAAAAGTACCAACCAACAAAGGTAACCTTAACCTTCAACAGTGGCAAGGA 1899  
QY 1801 ATCTAGTTTACTTAAGAAAGTGGCGAGCTTGTACTTAATGACAGAGTCTACTTGTCT 1860  
DB 1900 ATCTAGTTTACTTAAGAAAGTGGCGAGCTTGTACTTAATGACAGAGTCTACTTGTCT 1959

QY 1861 AGACGTAGAAAAGAAATGCT 1881  
Db 1960 AGACGTAGAAAAGAAATGCT 1980

RESULT 5  
BD193451

LOCUS BD193451 15614 bp DNA linear PAT 17-UTL-2003

DEFINITION Enterococcus faecialis polynucleotides and polypeptides.

ACCESSION BD193451

VERSION BD193451.1 GI:33003190

KEYWORDS JP 2002529046-A/45.

SOURCE unidentified

ORGANISM unidentified

REFERENCE 1 (bases 1 to 15614)  
Kunsch,C.A., Dillon,P.J. and Baraeh,S.C.  
Enterococcus faecialis polynucleotides and polypeptides  
Parent: JP 2002529046-A 45 03-SEP-2002;  
HUMAN GENOME SCIENCES INC

COMMENT OS Unidentified  
PN JP 2002529046-A/45  
PD 03-SEP-2002  
PF 04-MAY-1998 JP 1998548302  
PR 06-MAY-1997 US 60/044031,16-MAY-1997 US 60/046655 PR  
14-NOV-1997 US 60/066009  
PI CHARLES A KUNSCH,PATRICK J DILLON,STEVEN C BARASH PC  
CI2N15/31,C07K14/315,C07K16/12,C12O1/68  
CC Strandedness: Double;  
CC Topology: Linear;  
CC Enterococcus faecialis polynucleotides and polypeptides FH

Key Location/Qualifiers  
FT source 1..15614

FEATURES  
source 1..15614  
Location/Qualifiers  
1..15614  
/organism="unidentified"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:32644"

ORIGIN

Query Match 100.0%; Score 1881; DB 2; Length 15614;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1881; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAAGCAATTAATAAAGTTGGTATGACACCGTTAGTACCTGTTACTAATTTTGGCACTT 60  
Db 12164 ATGAAGCAATTAATAAAGTTGGTATGACACCGTTAGTACCTGTTACTAATTTTGGCACTT 12223

QY 61 TTCACAAGGTATTTAGGACAACAACGATTTGACAGAATAATGGGAGAGCGACAG 120  
Db 12224 TTCACAAGGTATTTAGGACAACAACGATTTGACAGAATAATGGGAGAGCGACAG 12283

QY 121 CTGCGATTACAAAAAGAAAAATGACGATTTTACGATCCGTTATTTCAAAATAGCGGG 180  
Db 12284 CTGCGATTACAAAAAGAAAAATGACGATTTTACGATCCGTTATTTCAAAATAGCGGG 12343

QY 181 AAAAGAAATGAGGAGTTGATTAATTCAGAGCTGGGAGATGTGACGTTTATTTAT 240  
Db 12344 AAAAGAAATGAGGAGTTGATTAATTCAGAGCTGGGAGATGTGACGTTTATTTAT 12403

QY 241 AACGTGACGAAGCAATTTTACGAGCAACGAGCGGCGCAAGCGTTGATGACGATAA 300  
Db 12404 AACGTGACGAAGCAATTTTACGAGCAACGAGCGGCGCAAGCGTTGATGACGATAA 12463

QY 301 CAAGCTGTCGAAGTTTAACTCTGCGGAAAACCTGTTGCTCAAGAAACAACGATGCAAT 360  
Db 12464 CAAGCTGTCGAAGTTTAACTCTGCGGAAAACCTGTTGCTCAAGAAACAACGATGCAAT 12523

QY 361 GGGATGTACTGTTAGTTACTATAAAAAAAGGTTAAAGTGCATGTTATACATT 420  
Db 12524 GGGATGTACTGTTAGTTACTATAAAAAAAGGTTAAAGTGCATGTTATACATT 12583

QY 421 AAAGAGAACCAAAAGAGGTGTAGTCTGCTACGAATATGCTGGCGTTCCAGTT 480  
Db 12584 AAAGAGAACCAAAAGAGGTGTAGTCTGCTACGAATATGCTGGCGTTCCAGTT 12643

QY 481 TACGAATATGATCAACCAACAGATGTTCTATAAATATGCAACGAAGAAATTACCGTT 540  
Db 12644 TACGAATATGATCAACCAACAGATGTTCTATAAATATGCAACGAAGAAATTACCGTT 12703

QY 541 GTTCAATTTTATCTTAAAAATGTGTAGCCAAATGATGTTATCATGTGAAAAAGTA 600  
Db 12704 GTTCAATTTTATCTTAAAAATGTGTAGCCAAATGATGTTATCATGTGAAAAAGTA 12763

QY 601 GGAATGCTGAAAAATGAAAGATTAATATGCGGAGATTGTTATTTCTTAAAAAGCAAGC 660  
Db 12764 GGAATGCTGAAAAATGAAAGATTAATATGCGGAGATTGTTATTTCTTAAAAAGCAAGC 12823

QY 661 TCACGAGCAAGTAAATATATCCAGAGAGTCAAAAGATGATTAATATCATGACCAACG 720  
Db 12824 TCACGAGCAAGTAAATATATCCAGAGAGTCAAAAGATGATTAATATCATGACCAACG 12883

QY 721 GATTAAGAAACAAGCAAAACGTTTATTACTGCGAAAAGTTATGAAATGGCGAAAATGAT 780  
Db 12884 GATTAAGAAACAAGCAAAACGTTTATTACTGCGAAAAGTTATGAAATGGCGAAAATGAT 12943

QY 781 TTCACAGAAAGCAGAAATGAAACGGGAAATTTAACGTTAAAAATCTTGAGGTTGTTG 840  
Db 12944 TTCACAGAAAGCAGAAATGAAACGGGAAATTTAACGTTAAAAATCTTGAGGTTGTTG 13003

QY 841 TATATTTTGAAGAGTAAAGCTCCAAATATGCAATTAATGAAATATGAAATCAACAAA 900  
Db 13004 TATATTTTGAAGAGTAAAGCTCCAAATATGCAATTAATGAAATATGAAATCAACAAA 13063

QY 901 ACACATTTTAACTTGAAGCAAAACAATCAACACCTGTTGAAAAACGTCAAAATGAT 960  
Db 13064 ACACATTTTAACTTGAAGCAAAACAATCAACACCTGTTGAAAAACGTCAAAATGAT 13123

QY 961 ACCTTAAAGTTGATTAACCAACCAACGATTTAGTGTAAAGATGTGGCAATTTGGCANA 1020  
Db 13124 ACCTTAAAGTTGATTAACCAACCAACGATTTAGTGTAAAGATGTGGCAATTTGGCANA 13183

QY 1021 AAAATTAATATCAAAATTTCTGTAAATATTCATTGGGGATTGCAACAAAGAGCGAC 1080  
Db 13184 AAAATTAATATCAAAATTTCTGTAAATATTCATTGGGGATTGCAACAAAGAGCGAC 13243

QY 1081 GCTAATAAATACGTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1140  
Db 13244 GCTAATAAATACGTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 13303

QY 1141 AACGTACTTCTGAGAGATATGCTTATGCTTATTAATTAATTAATTAATTAATTAAT 1200  
Db 13304 AACGTACTTCTGAGAGATATGCTTATGCTTATTAATTAATTAATTAATTAATTAAT 13363

QY 1201 GAAAATTAATCAAGTATGTAACACGAATATGCTTCACTGTGCGGCTTAACTTAACT 1260  
Db 13364 GAAAATTAATCAAGTATGTAACACGAATATGCTTCACTGTGCGGCTTAACTTAACT 13423

QY 1261 ATTCTAGCTTAACACGAGCGGCAACTTAAATTCGTTTACTTATATGCAATTAATGAA 1320  
Db 13424 ATTCTAGCTTAACACGAGCGGCAACTTAAATTCGTTTACTTATATGCAATTAATGAA 13483

QY 1321 AAAGCAGATCTTACGAAGAGCTTTAAAAATGAGCGAATGTTGATTAACGGTATACGAC 1380  
Db 13484 AAAGCAGATCTTACGAAGAGCTTTAAAAATGAGCGAATGTTGATTAACGGTATACGAC 13543

QY 1381 GACCAAAACCAACCAACTGTTGAAGTTGGAACAGGTGGGAAAACGTTTCAATTAATGCGAT 1440  
Db 13544 GACCAAAACCAACCAACTGTTGAAGTTGGAACAGGTGGGAAAACGTTTCAATTAATGCGAT 13603

QY 1441 GCGAGTGTGACGAGCAACAACCTTGGCGGAGCTTCTTGTGCTCGTGCATCAAAAC 1500  
Db 13604 GCGAGTGTGACGAGCAACAACCTTGGCGGAGCTTCTTGTGCTCGTGCATCAAAAC 13663

|    |       |                                                              |       |
|----|-------|--------------------------------------------------------------|-------|
| Qy | 1501  | AGCGACACAGAAATTATTTGAAATGTGATGAAACACGAAACAGCAACTTGGGTGAA     | 1560  |
| Db | 13664 | AGCGACACAGAAATTATTTGAAATGTGATGAAACACGAAACAGCAACTTGGGTGAA     | 13723 |
| Qy | 1561  | ACAAAGCTGAGCACTACTTTTACAAACAACGGCTGATGGATTAGTTGATATCACAGGG   | 1620  |
| Db | 13724 | ACAAAGCTGAGCACTACTTTTACAAACAACGGCTGATGGATTAGTTGATATCACAGGG   | 13783 |
| Qy | 1621  | CTTAATACGGTACCTATTTATTTAGAGAAACTGTAGCTTCGTGATGATTATGCTTGTGA  | 1680  |
| Db | 13784 | CTTAATACGGTACCTATTTATTTAGAGAAACTGTAGCTTCGTGATGATTATGCTTGTGA  | 13843 |
| Qy | 1681  | ACAAATCGGATGTAATTTGTGTCAATGAACAATCATATGCGCAACAGAAACCTTAGTT   | 1740  |
| Db | 13844 | ACAAATCGGATGTAATTTGTGTCAATGAACAATCATATGCGCAACAGAAACCTTAGTT   | 13903 |
| Qy | 1741  | TCACGAGAAAAAGTACCAACAAACAAAGTACCTTACCTTCAACAGGTGGCAAGGA      | 1800  |
| Db | 13904 | TCACGAGAAAAAGTACCAACAAACAAAGTACCTTACCTTCAACAGGTGGCAAGGA      | 13963 |
| Qy | 1801  | ATCTACGTTTACTTAGAGATGGCGCAGTCTTGCTACTTATTTGACAGAGTCTACTTTGCT | 1860  |
| Db | 13964 | ATCTACGTTTACTTAGAGATGGCGCAGTCTTGCTACTTATTTGACAGAGTCTACTTTGCT | 14023 |
| Qy | 1861  | AGACGTGAAAAAGAAATGCT                                         | 1881  |
| Db | 14024 | AGACGTGAAAAAGAAATGCT                                         | 14044 |

|            |                                       |         |     |        |
|------------|---------------------------------------|---------|-----|--------|
| RESULT 6   |                                       |         |     |        |
| CQ969453   |                                       |         |     |        |
| LOCUS      | CQ969453                              | 1875 bp | DNA | linear |
| DEFINITION | Sequence 28 from Patent WO2004106367. |         |     |        |
| ACCESSION  | CQ969453                              |         |     |        |
| VERSION    | CQ969453.1                            |         |     |        |
| KEYWORDS   | GI:56743584                           |         |     |        |
| SOURCE     |                                       |         |     |        |
| ORGANISM   | Enterococcus faecalis                 |         |     |        |
|            | Enterococcus faecalis                 |         |     |        |

| FEATURES | REFERENCE                                          |
|----------|----------------------------------------------------|
| source   | 1. 1875                                            |
|          | location/Qualifiers                                |
|          | InterCell AG (AT)                                  |
|          | Patent: WO 2004106567-A 28 09-DEC-2004;            |
|          | Enterococcus antigens                              |
|          | Meinkens, A., Nagy, E., Hammer, M. and Gelmann, D. |
|          | 1. 1875                                            |
|          | 1. 1875                                            |

```

/organism="Enterococcus faecalis"
/mol_type="unassigned DNA"
/db_xref="taxon:1351"

```

ORIGIN

|                            |       |              |       |              |
|----------------------------|-------|--------------|-------|--------------|
| Query Match                | 98.8% | Score 1859;  | DB 2; | Length 1875; |
| Best Local Similarity      | 99.7% | Pred. No. 0; |       |              |
| Matches 1875; Conservative | 0;    | Mismatches   | 6;    | Gaps 1;      |

[illegible]

|    |      |                                                                 |      |
|----|------|-----------------------------------------------------------------|------|
| Db | 241  | AACTGACGAAAGAAATTTTACGACGAAAGACGGACGGACGTAAGCTTGAATGACAGCTAAA   | 300  |
| OY | 301  | CAAGCTGTCGAAAGTTTAATCTCTGGGAAACCTGTGCTCAAGAACCAACCGATGCAAT      | 360  |
| Db | 301  | CMACTGTCCAAAGTTTAATCTCTGGGAAACCTGTGCTCAAGAACCAACCGATGCAAT       | 360  |
| OY | 361  | GGGAATGCTCACTGTTCACTTAACCTPAAAAAAAMAAATGGTAAGATGACGTGTATCAATT   | 420  |
| Db | 361  | GGGAATGCTCACTGTTCACTTAACCTPAAAAAAAMAAATGGTAAGATGACGTGTATCACTT   | 420  |
| OY | 421  | AAAGAAAGAACCAAAAGAGGGGTGATGTTGCTGTCTACGAATATGATGGTGGCCGTCCCACTT | 480  |
| Db | 421  | AAAGAAAGAACCAAAAGAGGGGTGATGTTGCTGTCTACGAATATGATGGTGGCCGTCCCACTT | 480  |
| OY | 481  | TACGAAATGATCAAGCAAAAGATGCTTCTATAAATATGGAACAGAAATTAAGCGGTT       | 540  |
| Db | 481  | TACGAAATGATCAAGCAAAAGATGCTTCTATAAATATGGAACAGAAATTAAGCGGTT       | 540  |
| OY | 541  | GTTCAATATTTATCTTAAAAATGTGGTAGCCAAATGATGGTACTTTCATGTGAAAAAAGTA   | 600  |
| Db | 541  | GTTCAATATTTATCTTAAAAATGTGGTAGCCAAATGATGGTACTTTCATGTGAAAAAAGTA   | 600  |
| OY | 601  | GGAATGCTGAAAAATGAAGATTAAATGGCCAGAAATTTGTTATTTCTTAAAGGGAAGGC     | 660  |
| Db | 601  | GGAATGCTGAAAAATGAAGATTAAATGGCCAGAAATTTGTTATTTCTTAAAGGGAAGGC     | 660  |
| OY | 651  | TCACAGGCAAGTAAATATATCAAGGAATCAAGATGGAATTAATATACATGACGACAAG      | 720  |
| Db | 651  | TCACAGGCAAGTAAATATATCAAGGAATCAAGATGGAATTAATATACATGACGACAAG      | 720  |
| OY | 721  | GATTAAGAACAGCAAAAACGCTTTATTACTGGGAAAAAGTTATGAAATTTGGCGAAATGAT   | 780  |
| Db | 721  | GATTAAGAACAGCAAAAACGCTTTATTACTGGGAAAAAGTTATGAAATTTGGCGAAATGAT   | 780  |
| OY | 775  | TTCAAGAGACAGAAATGGAACGGGAAATTAACAGTTAAAAATCTTGAAGTTGGTTCG       | 840  |
| Db | 775  | TTCAAGAGACAGAAATGGAACGGGAAATTAACAGTTAAAAATCTTGAAGTTGGTTCG       | 840  |
| OY | 841  | TATATTTTAAAGAAAGTAAAGCTCCAAATTAATGCAAAATTAATGAAATCAACAAATA      | 900  |
| Db | 841  | TATATTTTAAAGAAAGTAAAGCTCCAAATTAATGCAAAATTAATGAAATCAACAAATA      | 900  |
| OY | 891  | ACACCATTTACATTTGAAGCAAAACATCAAAACACTGTGTGAAAAAAACAGTCAAAAATGAT  | 960  |
| Db | 891  | ACACCATTTACATTTGAAGCAAAACATCAAAACACTGTGTGAAAAAAACAGTCAAAAATGAT  | 960  |
| OY | 961  | ACCTCTAAAGTTGATAAACAAACCAAGCTTAGATGCTAAAGATGTGGCAATTTGGCGAA     | 1020 |
| Db | 961  | ACCTCTAAAGTTGATAAACAAACCAAGCTTAGATGCTAAAGATGTGGCAATTTGGCGAA     | 1020 |
| OY | 1021 | AAAAATTAATATCAAAATTTCTGTAAATATTCATTTGGGATTTGCAGACAAAGAAAGCGAC   | 1080 |
| Db | 1021 | AAAAATTAATATCAAAATTTCTGTAAATATTCATTTGGGATTTGCAGACAAAGAAAGCGAC   | 1080 |
| OY | 1081 | GCTATATAATATGCTCAAAATTTCAATTTAGTTGATTAACATGATGACGCTTAACTTTGAT   | 1140 |
| Db | 1081 | GCTATATAATATGCTCAAAATTTCAATTTAGTTGATTAACATGATGACGCTTAACTTTGAT   | 1140 |
| OY | 1141 | AACGTGACTTGGAGGATATGCTTATGCGTTATATGATGGGATATACGATGATGCTCCT      | 1200 |
| Db | 1141 | AACGTGACTTGGAGGATATGCTTATGCGTTATATGATGGGATATACGATGATGCTCCT      | 1200 |
| OY | 1201 | GAAAAATTAATCAAGATGCTGAACAGCAATGGCTTCACTGTGCGCGCTTAAATCCAGCGAT   | 1260 |
| Db | 1201 | GAAAAATTAATCAAGATGCTGAACAGCAATGGCTTCACTGTGCGCGCTTAAATCCAGCGAT   | 1260 |
| OY | 1261 | ATTCTCAAGCTAACACGAGGCGGACACTAAATTTGTTACTTATGCAATTTAAATGAA       | 1320 |
| Db | 1261 | ATTCTCAAGCTAACACGAGGCGGACACTAAATTTGTTACTTATGCAATTTAAATGAA       | 1320 |
| OY | 1321 | AAACGATCTCTACGAAAGGCTTTAAAAATGAGCGAATGTTGATTAACGTCATACCGAC      | 1380 |

Db 1315 AAAGCAGATCTACGAAAGCTTTAAATAATGAGCGAATGTTGATPAAGGTCATACCGAC 1374  
Qy 1381 GACCAAAACCAACCAACTGTTGAAGTTTGACAGGTGGGAAACGTTTCACTTAAAGTCAT 1440  
Db 1375 GACCAAAACCAACCAACTGTTGAAGTTTGACAGGTGGGAAACGTTTCACTTAAAGTCAT 1434  
Qy 1441 GCGCATGTGACAGCGACACAGCCTTGGCGGAGCTTCCCTTGTGCTCGGTATCAAAAC 1500  
Db 1435 GCGCATGTGACAGCGACACAGCCTTGGCGGAGCTTCCCTTGTGCTCGGTATCAAAAC 1494  
Qy 1501 AGCGACACAGCAATTTATTTGAAATCGATGAAACAGCAAGACAGCAACTTGGGTGAA 1560  
Db 1495 AGCGACACAGCAATTTATTTGAAATCGATGAAACAGCAAGACAGCAACTTGGGTGAA 1554  
Qy 1561 ACAAAGCTGAGCACTCTTTTACACACAGGCTGATGATTTAGTTATATCAACAGG 1620  
Db 1555 ACAAAGCTGAGCACTCTTTTACACACAGGCTGATGATTTAGTTATATCAACAGG 1614  
Qy 1621 CTTAATAACGTAATCTATTTAGAAAGCTGTAGCTCGATGATTTATGCTTGT 1680  
Db 1615 CTTAATAACGTAATCTATTTAGAAAGCTGTAGCTCGATGATTTATGCTTGT 1674  
Qy 1681 ACAAATCGGATTTGAAATTTGTGCTCAATGAAATCATATGCAACAGCAAACTTGT 1740  
Db 1675 ACAAATCGGATTTGAAATTTGTGCTCAATGAAATCATATGCAACAGCAAACTTGT 1734  
Qy 1741 TCACAGAAAGAAAGTACCAACCAAGGTAACCTTCACTTCAACAGGTGGCAAG 1800  
Db 1735 TCACAGAAAGAAAGTACCAACCAAGGTAACCTTCACTTCAACAGGTGGCAAG 1794  
Qy 1801 ATCTAGTTTACTTGAAGTGGCGAGCTTGTCTACTTATTCAGAGTCTTACTTGT 1860  
Db 1795 ATCTAGTTTACTTGAAGTGGCGAGCTTGTCTACTTATTCAGAGTCTTACTTGT 1854  
Qy 1861 AGACGTGAGAAAGAAATGCT 1881  
Db 1855 AGACGTGAGAAAGAAATGCT 1875

RESULT 7  
AE016830\_10  
WPCOMMENT

Sequence split into 33 fragments LOCUS AE016830 Accession AE016830  
Fragment Name Begin End  
AE016830\_01 1 110000  
AE016830\_02 100001 210000  
AE016830\_03 200001 310000  
AE016830\_04 300001 410000  
AE016830\_05 400001 510000  
AE016830\_06 500001 610000  
AE016830\_07 600001 710000  
AE016830\_08 700001 810000  
AE016830\_09 800001 910000  
AE016830\_10 900001 1010000  
AE016830\_11 1000001 1110000  
AE016830\_12 1100001 1210000  
AE016830\_13 1200001 1310000  
AE016830\_14 1300001 1410000  
AE016830\_15 1400001 1510000  
AE016830\_16 1500001 1610000  
AE016830\_17 1600001 1710000  
AE016830\_18 1700001 1810000  
AE016830\_19 1800001 1910000  
AE016830\_20 1900001 2010000  
AE016830\_21 2000001 2110000  
AE016830\_22 2100001 2210000  
AE016830\_23 2200001 2310000  
AE016830\_24 2300001 2410000  
AE016830\_25 2400001 2510000  
AE016830\_26 2500001 2610000  
AE016830\_27 2600001 2710000  
AE016830\_28 2700001 2810000  
AE016830\_29 2800001 2910000

AE016830\_29 2900001 3010000  
AE016830\_30 3000001 3110000  
AE016830\_31 3100001 3210000  
AE016830\_32 3200001 3218031  
Continuation (11 of 33) of AE016830 from base 1000001 (AE016830 Enterococcus faecalis V5  
Query Match 98.8%; Score 1859; DB 15; Length 110000;  
Best Local Similarity 99.7%; Pred. No. 0;  
Matches 1875; Conservative 0; Mismatches 0; Indels 6; Gaps 1;  
Qy 1 ATGAAGCAATTAATAAAGTTTGGTACACCGTTAGTCTTGTATCTAATTTTGGCACTT 60  
Db 62129 ATGAAGCAATTAATAAAGTTTGGTACACCGTTAGTCTTGTATCTAATTTTGGCACTT 62188  
Qy 61 TTCACAGTGTATTTGGGACCAACCTGATTTTGGAGAAAGGGGAGGCCACAG 120  
Db 62189 TTCACAGTGTATTTGGGACCAACCTGATTTTGGAGAAAGGGGAGGCCACAG 62248  
Qy 121 CTGCTGATTCAAAAAGAAATGACGATTTTACAGATCCGCTTATTCAAAATAGCGG 180  
Db 62249 CTGCTGATTCAAAAAGAAATGACGATTTTACAGATCCGCTTATTCAAAATAGCGG 62308  
Qy 181 AAAGAAATGACGAGTTGTGATTAATATCAAGACTGGCAGATGTGACGTTTAGTAT 240  
Db 62309 AAAGAAATGACGAGTTGTGATTAATATCAAGACTGGCAGATGTGACGTTTAGTAT 62368  
Qy 241 AACGTGACGAAGATTTTACAGACAGAGCGGACAGGGGTTGATGACACTAA 300  
Db 62369 AACGTGACGAAGATTTTACAGACAGAGCGGACAGGGGTTGATGACACTAA 62428  
Qy 301 CAAGCTGTCCAAAGTTTAACTCTCGGAAACCTGTGTCTCAAGGAACCAACGATGCAAT 360  
Db 62429 CAAGCTGTCCAAAGTTTAACTCTCGGAAACCTGTGTCTCAAGGAACCAACGATGCAAT 62488  
Qy 361 GGGAAATGTCACTGTTCACTTACTTAAAAACAAATGTTAAAGATGCACTATACATT 420  
Db 62489 GGGAAATGTCACTGTTCACTTACTTAAAAACAAATGTTAAAGATGCACTATACATT 62548  
Qy 421 AAAGAAAGCAAAAAGAGGCTGTAGTGTCTGACGAATATGTTGGGCGTTCCAGTT 480  
Db 62549 AAAGAAAGCAAAAAGAGGCTGTAGTGTCTGACGAATATGTTGGGCGTTCCAGTT 62608  
Qy 481 TACGAATGATCAACCAACAGATGTTCTTAATAATGAAACAGAAATTAAGCGTT 540  
Db 62609 TACGAATGATCAACCAACAGATGTTCTTAATAATGAAACAGAAATTAAGCGTT 62668  
Qy 541 GTTCAATTTATCTTAATAATGTTGTTAGTATGATGTTTCAATGAAAAAGTA 600  
Db 62669 GTTCAATTTATCTTAATAATGTTGTTAGTATGATGTTTCAATGAAAAAGTA 62722  
Qy 601 GGAATCGCTGAAGATGAAGATTAAATGGCGAGAAATTTGTTATTTCTTAAAGCGAAGC 660  
Db 62723 GGAATCGCTGAAGATGAAGATTAAATGGCGAGAAATTTGTTATTTCTTAAAGCGAAGC 62782  
Qy 661 TCACAGGCAAGTAATATATCAAGAGTCAAGATGATTAATATCATGACCAACG 720  
Db 62783 TCACAGGCAAGTAATATATCAAGAGTCAAGATGATTAATATCATGACCAACG 62842  
Qy 721 GATTAAGAACAGCAAAACGCTTATTAATCTGGGAAAAAGTTAGAAATGGCGAAAAATGAT 780  
Db 62843 GATTAAGAACAGCAAAACGCTTATTAATCTGGGAAAAAGTTAGAAATGGCGAAAAATGAT 62902  
Qy 781 TTCACAGAAAGCAGAAATGAGACGGAGAAATTAACAGTTTAAATAATCTTGAGGTTGGT 840  
Db 62903 TTCACAGAAAGCAGAAATGAGACGGAGAAATTAACAGTTTAAATAATCTTGAGGTTGGT 62962  
Qy 841 TATATTTTGAAGAGTAAGCTCCAAATATGCAATTAATTAATGAAAAATCAACAAAA 900  
Db 62963 TATATTTTGAAGAGTAAGCTCCAAATATGCAATTAATTAATGAAAAATCAACAAAA 63022  
Qy 901 ACACATTTAACAATGAAGCAACAACTCAACCTGTTGAAAAAAGTCAAAAATGAT 960  
Db 63023 ACACATTTAACAATGAAGCAACAACTCAACCTGTTGAAAAAAGTCAAAAATGAT 63082

```

QY 961 ACCTCAAAAGTTGATPAAACAAACCAAGCTTAGATGTAAAGATGTGGCAATTGGCCGA 1020
Db 63083 ACCCTTAAAGTTGATPAAACAAACCAAGCTTAGATGTAAAGATGTGGCAATTGGCCGA 63142
QY 1021 AAAATTAATATCAAAATTTCTGTAAATATTTCCATTGGGGATTCAGACAAAGAGCGAC 1080
Db 63143 AAAATTAATATCAAAATTTCTGTAAATATTTCCATTGGGGATTCAGACAAAGAGCGAC 63202
QY 1081 GCTAATTAATACGTCAAAATTCATTATGTTGATPAAACATGATGCAGCCTTAACTTTTGT 1140
Db 63203 GCTAATTAATACGTCAAAATTCATTATGTTGATPAAACATGATGCAGCCTTAACTTTTGT 63262
QY 1141 AACGGAAGCTTCGGAGAGATATCTTATGCTTATATGATGGGGATPACATGATGCTCT 1200
Db 63263 AACGGAAGCTTCGGAGAGATATCTTATGCTTATATGATGGGGATPACATGATGCTCT 63322
QY 1201 GAAATTAATCAAGTGAAGTGAACAGCAATGGCTTCACTGCGCCGTTAATCAGCGTAT 1260
Db 63323 GAAATTAATCAAGTGAAGTGAACAGCAATGGCTTCACTGCGCCGTTAATCAGCGTAT 63382
QY 1261 ATTCTAGCGTTAACACAGGCGGCACTPAAATTCGTTTACTTTATGCAATTTAAATGA 1320
Db 63383 ATTCTAGCGTTAACACAGGCGGCACTPAAATTCGTTTACTTTATGCAATTTAAATGA 63442
QY 1321 AAAGAGATCCCTACGAAAGGCTTTAAATAATGAGCGGAATGTTGATPACGGTCAATCCGAC 1380
Db 63443 AAAGAGATCCCTACGAAAGGCTTTAAATAATGAGCGGAATGTTGATPACGGTCAATCCGAC 63502
QY 1381 GACCAAAACACCACTGTTGAAGTGTGACAGGTGGGAAACGTTTCAATTAAGTCGAT 1440
Db 63503 GACCAAAACACCACTGTTGAAGTGTGACAGGTGGGAAACGTTTCAATTAAGTCGAT 63562
QY 1441 GCGGATGTGACAGCGACACAGACCTTTGGCGGAGCTTCTTTGTCGTCGATCAAAAC 1500
Db 63563 GCGGATGTGACAGCGACACAGACCTTTGGCGGAGCTTCTTTGTCGTCGATCAAAAC 63622
QY 1501 AGCGACACAGCAATTTATTTGAAATTCGATGAAACAGCAAGCGCACTTGGGTGAA 1560
Db 63623 AGCGACACAGCAATTTATTTGAAATTCGATGAAACAGCAAGCGCACTTGGGTGAA 63682
QY 1561 ACAAAAGCTGAAGCACTCTTTTCAACAACGGCTGATGATTAAGTTATACAGGG 1620
Db 63683 ACAAAAGCTGAAGCACTCTTTTCAACAACGGCTGATGATTAAGTTATACAGGG 63742
QY 1621 CTTAAATACGTAAGCTTATTTAGAGAACTGTAGCTCTGATGATTAAGTTATGTTT 1680
Db 63743 CTTAAATACGTAAGCTTATTTAGAGAACTGTAGCTCTGATGATTAAGTTATGTTT 63802
QY 1681 ACAATTCGATGATTTGTGTGATGAAACATCATATGAGCAACAGAAACCTAGTT 1740
Db 63803 ACAATTCGATGATTTGTGTGATGAAACATCATATGAGCAACAGAAACCTAGTT 63862
QY 1741 TCACCAAGAAAAGTACCAACCAACAGAAAGTACCTTACCTTCAACAGGTGGCAAGGA 1800
Db 63863 TCACCAAGAAAAGTACCAACCAACAGAAAGTACCTTACCTTCAACAGGTGGCAAGGA 63922
QY 1801 ATCTACGTTTACTTGAAGAGTGGCGCAGCTTGTCTACTTATTTGCAAGAGTCTACTTGT 1860
Db 63923 ATCTACGTTTACTTGAAGAGTGGCGCAGCTTGTCTACTTATTTGCAAGAGTCTACTTGT 63982
QY 1861 AGACCTAGAAAAGAAATGCT 1881
Db 63983 AGACCTAGAAAAGAAATGCT 64003

```

```

RESULT 8
LOCUS BD191808 1687 bp DNA linear part 17-Jul-2003
DEFINITION Enterococcus Faecalis polynucleotides and polypeptides.
ACCESSION BD191808
VERSION BD191808.1 GI:33001547
KEYWORDS JP 2002516571-A/110.

```

```

SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE 1 (bases 1 to 1687)
AUTHORS Kunsch, C.A., Choi, G.H., Bailey, C. and Hromockyj, A.
TITLE Enterococcus Faecalis polynucleotides and polypeptides
JOURNAL Patent: JP 2002516571-A 110 04-JUN-2002;
HUMAN GENOME SCIENCES INC
COMMENT
PN JP 2002516571-A/110
PD 04-JUN-2002
PF 04-MAY-1998 JP 1998548289
PR 06-MAY-1997 US 60/044031, 16-MAY-1997 US 60/046655 PR
PI 14-NOV-1997 US 60/066009
PI CHARLES A KUNSCH, GIL H CHOI, CAMELIA BAILEY, ALEX HROMOCKYJ PC
CI 21N15/31, C07K14/315, C07K16/12, C12Q1/68, C12N1/21, C12N5/12, PC
G01N33/569,
PC G01N33/68, A61K39/09
CC Scrandedness: Double;
CC Topology: Linear;
FH Key location/Qualifiers.

FEATURES
source location/Qualifiers
1..1687
/organism="synthetic construct"
/mol type="genomic DNA"
/db_xref="taxon:32630"

ORIGIN
Query Match 89.7%; Score 1687; DB 2; Length 1687;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1687; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 96 AAAAAAAAAAGGAGAGAGGCAAGCTGCTGATTCACAAAAGAAAATGACGATTTAC 155
Db 1 AAAAAAAAAAGGAGAGAGGCAAGCTGCTGATTCACAAAAGAAAATGACGATTTAC 60
QY 156 AGATCCGCTTATTTCAAAATATGCGGAAAGAAATGAGCGAGTTGATTAATATCAAGACT 215
Db 61 AGATCCGCTTATTTCAAAATATGCGGAAAGAAATGAGCGAGTTGATTAATATCAAGACT 120
QY 216 GGCAGATGTGAGTTAGTATTATTAATACGTGACGAACGAATTTTACGACCAACGAGCGC 275
Db 121 GGCAGATGTGAGTTAGTATTATTAATACGTGACGAACGAATTTTACGACCAACGAGCGC 180
QY 276 AGCGCAAGCGTTGATGACGCTAAACAGCTGTCCAAAGTTTAACTCCCTGGGAACTGT 335
Db 181 AGCGCAAGCGTTGATGACGCTAAACAGCTGTCCAAAGTTTAACTCCCTGGGAACTGT 240
QY 336 TGCTCAAGAACCAACGATGCAATGCGAATGTCTCACTGTTCACTTAACCTAATAAACA 395
Db 241 TGCTCAAGAACCAACGATGCAATGCGAATGTCTCACTGTTCACTTAACCTAATAAACA 300
QY 396 TGTAAAGATGAGTGTATACCATTAATAAGAAACCAAAAGAGGTGATGTTGCTGTAC 455
Db 301 TGTAAAGATGAGTGTATACCATTAATAAGAAACCAAAAGAGGTGATGTTGCTGTAC 360
QY 456 GAATATGTGTGTGGGTTCCCGTTTACGAATGATCAAGCAACAGATGCTTCTATTA 515
Db 361 GAATATGTGTGTGGGTTCCCGTTTACGAATGATCAAGCAACAGATGCTTCTATTA 420
QY 516 ATATGAAACAGAAATATAGCGTTGTTCAATTAATCTTAATAATGTGTATGCAATGA 575
Db 421 ATATGAAACAGAAATATAGCGTTGTTCAATTAATCTTAATAATGTGTATGCAATGA 480
QY 576 TGTATGTTTACATGTGAATAAAGTGAATCTGCTGAATAATGAAGATTAATGCGCGCA 635
Db 481 TGTATGTTTACATGTGAATAAAGTGAATCTGCTGAATAATGAAGATTAATGCGCGCA 540
QY 636 ATTGTTATTTCTAATAAGCAAGGCTCCAGGCAAGTAAATATATCCAGAGAGTCA 695
Db 541 ATTGTTATTTCTAATAAGCAAGGCTCCAGGCAAGTAAATATATCCAGAGAGTCA 600
QY 696 AGATGATTAATATACATGACCAACGATPAAAGAACCAACCAACGCTTATTAATCTGGGA 755

```

DB 601 AGATGATTATATACATGGACAACGGATTAAGAACAGCAAAACGTTTATTACTGGAA 660  
QY 756 AAGTTATGAAATTTGGCGAAATGATTTTCACAGACGAGAAATGGAACGGGAATTAC 815  
DB 661 AAGTTATGAAATTTGGCGAAATGATTTTCACAGACGAGAAATGGAACGGGAATTAC 720  
QY 816 AGTTAAAAATCTTGAAGTTGGTTCGTATATTTTAGAAGAAAGTAAAGCTCCAAATATAC 875  
DB 721 AGTTAAAAATCTTGAAGTTGGTTCGTATATTTTAGAAGAAAGTAAAGCTCCAAATATAC 780  
QY 876 AGAATTAATTTGAAAAATCAAAACAAAAACACATTTCATTAATGAAGCAAAACAAACACC 935  
DB 781 AGAATTAATTTGAAAAATCAAAACAAAAACACATTTCATTAATGAAGCAAAACAAACACC 840  
QY 936 TGTGAAAAAATCAAGTCAAAAAATGATACCTCTAAAGTTGATTAACCAACCAAGCTTAA 995  
DB 841 TGTGAAAAAATCAAGTCAAAAAATGATACCTCTAAAGTTGATTAACCAACCAAGCTTAA 900  
QY 996 TGTGAAAAATGTCGCAATTTGGCAAAATTTAAATATCAATTTTCTGTAATATTTCCATT 1055  
DB 901 TGTGAAAAATGTCGCAATTTGGCAAAATTTAAATATCAATTTTCTGTAATATTTCCATT 960  
QY 1056 GGGGATTTGACACAAAGAAAGCGACTAATAATAACGTCAAAATTCATTTAGTTGATA 1115  
DB 961 GGGGATTTGACACAAAGAAAGCGACTAATAATAACGTCAAAATTCATTTAGTTGATA 1020  
QY 1116 ACATGATGACGCTTAACTTTGATTAACGTGACTTCTGAGAGTATGCTTAATGCTTAA 1175  
DB 1021 ACATGATGACGCTTAACTTTGATTAACGTGACTTCTGAGAGTATGCTTAATGCTTAA 1080  
QY 1176 TGAATGGGATACAGTATGCTCTGAAAAATTCAGAGTGAACAGCAAAATGAGCTT 1235  
DB 1081 TGAATGGGATACAGTATGCTCTGAAAAATTCAGAGTGAACAGCAAAATGAGCTT 1140  
QY 1236 CACTGTCGCGCTTAATCAGCGTATATTCCTACGCTAACACAGCGGCGACACTAAAT 1295  
DB 1141 CACTGTCGCGCTTAATCAGCGTATATTCCTACGCTAACACAGCGGCGACACTAAAT 1200  
QY 1296 CGTTTAACTTATGCAATTTAAATGAAGAAAGCAGATCCTACGAAGGCTTTAAATGAAGC 1355  
DB 1201 CGTTTAACTTATGCAATTTAAATGAAGAAAGCAGATCCTACGAAGGCTTTAAATGAAGC 1260  
QY 1356 GAATGTTGATAACGTCATACCGACGACCAACCAACCACTGTTGAAGTTGTGACAG 1415  
DB 1261 GAATGTTGATAACGTCATACCGACGACCAACCAACCACTGTTGAAGTTGTGACAG 1320  
QY 1416 TGGGAAACGTTTCATTAAGTCGATGCGATGTGACAGCGACACAGCTTGGCGGAGC 1475  
DB 1321 TGGGAAACGTTTCATTAAGTCGATGCGATGTGACAGCGACACAGCTTGGCGGAGC 1380  
QY 1476 TTCCCTTTCGTCGCGATCAAAACAGCGACACAGCAATTAATTTGAAATGATGAAC 1535  
DB 1381 TTCCCTTTCGTCGCGATCAAAACAGCGACACAGCAATTAATTTGAAATGATGAAC 1440  
QY 1536 AACGAAGCAGCACTTGGGTGAAAAACAAAACCTGAGCACTACTTTTACCAACAGC 1595  
DB 1441 AACGAAGCAGCACTTGGGTGAAAAACAAAACCTGAGCACTACTTTTACCAACAGC 1500  
QY 1596 TGAATGATTTAGTTGATATACAGGGCTTAATACGCTTATTTTGAAGAAACCTGT 1655  
DB 1501 TGAATGATTTAGTTGATATACAGGGCTTAATACGCTTATTTTGAAGAAACCTGT 1560  
QY 1656 AGCTCTGATGATTTATGCTTTTAAACAAATCGGATTTGATTTGATTTGATTTGATTT 1715  
DB 1561 AGCTCTGATGATTTATGCTTTTAAACAAATCGGATTTGATTTGATTTGATTTGATTT 1620  
QY 1716 ATATGCAACAAGAAAACTTAGTTTACACAGAAAAAGTACCAACCAACCAAGGTAAC 1775  
DB 1621 ATATGCAACAAGAAAACTTAGTTTACACAGAAAAAGTACCAACCAACCAAGGTAAC 1680  
QY 1776 CTTAACCT 1782  
DB 1681 CTTAACCT 1687

RESULT 9  
LOCUS AR228553 1687 bp DNA linear PAT 20-DEC-2002  
DEFINITION Sequence 219 from patent US 6448043.  
ACCESSION AR228553  
VERSION AR228553.1 GI:27267390  
KEYWORDS  
SOURCE  
ORGANISM Unknown.  
REFERENCE  
1 (bases 1 to 1687)  
Choi, G.-H., Bailey, C., Hromockyj, A. and Kunsch, C.A.  
Enterococcus faecalis EF040 and uses therefor  
Patent: US 6448043-A 219 10-SEP-2002;  
JOURNAL Human Genome Sciences, Inc.; Rockville, MD;  
WOX;  
FEATURES  
location/Qualifiers  
source 1..1687  
/organism="unknown"  
/mol\_type="genomic DNA"  
ORIGIN  
Query Match 89.7%; Score 1687; DB 2; Length 1687;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1687; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 96 AGAAGAAAAATGGGGAGAGCGCACAGCTCGTGATTCACAAAAAGAAATGACGGATTAC 155  
DB 1 AGAAGAAAAATGGGGAGAGCGCACAGCTCGTGATTCACAAAAAGAAATGACGGATTAC 60  
QY 156 AGATCGGCTTATTTCAAAATTAAGCGGGAAGAAATGAGAGTTTGAATAATTAAGAGACT 215  
DB 61 AGATCGGCTTATTTCAAAATTAAGCGGGAAGAAATGAGAGTTTGAATAATTAAGAGACT 120  
QY 216 GGCAGATGTGACGTTAGTATTTATTAACGTGACGACCAATTTTACGACACAGACGCGC 275  
DB 121 GGCAGATGTGACGTTAGTATTTATTAACGTGACGACCAATTTTACGACACAGACGCGC 180  
QY 276 AGGCGCAAGCGTTGATGACGCTTAAACAGCTGTCCAAAGTTTAACTCCTGGGAAACCTGT 335  
DB 181 AGGCGCAAGCGTTGATGACGCTTAAACAGCTGTCCAAAGTTTAACTCCTGGGAAACCTGT 240  
QY 336 TGTCTAAGGAACCAACGATGCAAAATGGGAATGTCTGCTGCTGATTAACCTAAACCAAA 395  
DB 241 TGTCTAAGGAACCAACGATGCAAAATGGGAATGTCTGCTGCTGATTAACCTAAACCAAA 300  
QY 396 TGTGTAAGATGACGATGTATACCATTAAGAAAGAACCAAAAGAGGTGATGTGCTAC 455  
DB 301 TGTGTAAGATGACGATGTATACCATTAAGAAAGAACCAAAAGAGGTGATGTGCTAC 360  
QY 456 GAATATGATGTGCGCTTCCAGTTTACGAAATGATCAAGCAACAGATGTTCTTATTA 515  
DB 361 GAATATGATGTGCGCTTCCAGTTTACGAAATGATCAAGCAACAGATGTTCTTATTA 420  
QY 516 ATATGGAACAGAGAAATTAAGGGTGTTCATTTATCTTAATAATGTGATGCAATGA 575  
DB 421 ATATGGAACAGAGAAATTAAGGGTGTTCATTTATCTTAATAATGTGATGCAATGA 480  
QY 576 TGTGATTTTACATGTGAAAAAAGTAAAGCTGTAATAATGAAGATTAATGAGCGCAGA 635  
DB 481 TGTGATTTTACATGTGAAAAAAGTAAAGCTGTAATAATGAAGATTAATGAGCGCAGA 540  
QY 636 ATTGTTATTTTCTTAAAGCGAAAGCTCACACAGGCAAGTAAATATTCACAGAGTCAA 695  
DB 541 ATTGTTATTTTCTTAAAGCGAAAGCTCACACAGGCAAGTAAATATTCACAGAGTCAA 600  
QY 696 AGATGATTTATTAATGAGCAACAGGATTAAGAAACAGAGAAACGCTTATTAATCTGGGA 755  
DB 601 AGATGATTTATTAATGAGCAACAGGATTAAGAAACAGAGAAACGCTTATTAATCTGGGA 660  
QY 756 AAGTTATGAAATTTGGCGAAATGATTTTCACAGACGAGAAATGGAACGGGAATTAC 815

Db 661 AAGTTAAGAAATGGCGAATAATGATTTACAGAAAGCAGAGATGGAACGGGAGAAATTAAC 720  
Qy 816 AGTTAAATCTGAAGTTGGTTCGTATATTTTGAAGAGTAAGTAAAGCTCAATAATATGC 875  
Db 721 AGTTAAATCTGAAGTTGGTTCGTATATTTTGAAGAGTAAGTAAAGCTCAATAATATGC 780  
Qy 876 AGAATTAATGAATAATCAACAAACACCAATTAACAATTGAAGCAACATCAACAC 935  
Db 781 AGAATTAATGAATAATCAACAAACACCAATTAACAATTGAAGCAACATCAACAC 840  
Qy 936 TGTGAAAAAAGCTCAAAATGATACCTTAAAGTTGATTAACACCAACGCTTGA 995  
Db 841 TGTGAAAAAAGCTCAAAATGATACCTTAAAGTTGATTAACACCAACGCTTGA 900  
Qy 996 TGTGAAAAAAGCTCAAAATGATACCTTAAAGTTGATTAACACCAACGCTTGA 1055  
Db 901 TGTGAAAAAAGCTCAAAATGATACCTTAAAGTTGATTAACACCAACGCTTGA 960  
Qy 1056 GGGGATTCAGACAAAGAGGCGAGCTAATTAATAGTCAAAATTAAGTTGATTA 1115  
Db 961 GGGGATTCAGACAAAGAGGCGAGCTAATTAATAGTCAAAATTAAGTTGATTA 1020  
Qy 1116 ACATGATGACGCTTAACCTTTGATTAACGTAACCTTGAAGATGCTTATGCTTATA 1175  
Db 1021 ACATGATGACGCTTAACCTTTGATTAACGTAACCTTGAAGATGCTTATGCTTATA 1080  
Qy 1176 TGAATGGGATGACGATATGCTTCTGTAATTAATCAAGTACGTAACGTAATGCTT 1235  
Db 1081 TGAATGGGATGACGATATGCTTCTGTAATTAATCAAGTACGTAACGTAATGCTT 1140  
Qy 1236 CACTGCGCGCTTAATCCAGCGTATATCTTACGCTTAACGCGCGGCACTTAAAT 1295  
Db 1141 CACTGCGCGCTTAATCCAGCGTATATCTTACGCTTAACGCGCGGCACTTAAAT 1200  
Qy 1296 CGTTTACTTAATGATTAATGAATAAGCAATCCTAAGAAAGCTTTAAATGAGG 1355  
Db 1201 CGTTTACTTAATGATTAATGAATAAGCAATCCTAAGAAAGCTTTAAATGAGG 1260  
Qy 1356 GAAATGTTAATGCGTATACCGACGACCAACCAACCACTGTTGAAGTTGTGACG 1415  
Db 1261 GAAATGTTAATGCGTATACCGACGACCAACCAACCACTGTTGAAGTTGTGACG 1320  
Qy 1416 TGGGAAACGTTTCAATTAAGTGAATGCGATGCGATGACGAGGACCAAGCTTGGCGGAGC 1475  
Db 1321 TGGGAAACGTTTCAATTAAGTGAATGCGATGCGATGACGAGGACCAAGCTTGGCGGAGC 1380  
Qy 1476 TTCTTTTGTGTCCTGATCAAAACAGCGACACAGCAATTAATTTGAATGATGAAC 1535  
Db 1381 TTCTTTTGTGTCCTGATCAAAACAGCGACACAGCAATTAATTTGAATGATGAAC 1440  
Qy 1536 AAGCAAGACGACCTTGGGTTGAAAACAAAGCTGAAGCACTACTTTTCAACACGCG 1595  
Db 1441 AAGCAAGACGACCTTGGGTTGAAAACAAAGCTGAAGCACTACTTTTCAACACGCG 1500  
Qy 1596 TGAATGATTAATGATTAATGAGGCTTAATAGCTTAAATTTGAAGAACTGT 1655  
Db 1501 TGAATGATTAATGATTAATGAGGCTTAATAGCTTAAATTTGAAGAACTGT 1560  
Qy 1656 AGCTCTGATGATTAATGCTTGTTAACAATCGATGTAATTTGCTCAATGAACAATC 1715  
Db 1561 AGCTCTGATGATTAATGCTTGTTAACAATCGATGTAATTTGCTCAATGAACAATC 1620  
Qy 1716 AATATGACACAAAGAAACCTTGTTCACCAAGAAAGTACCAACCAACCAAGAGTAC 1775  
Db 1621 AATATGACACAAAGAAACCTTGTTCACCAAGAAAGTACCAACCAACCAAGAGTAC 1680  
Qy 1776 CTTACCT 1782  
Db 1681 CTTACCT 1687

RESULT 10

AR692968  
LOCUS AR692968 1687 bp DNA 1linear PAT 14-SEP-2005  
DEFINITION Sequence 219 from patent US 6913907.  
ACCESSION AR692968  
VERSION AR692968.1 GI:75182711  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE  
1 (bases 1 to 1687)  
Choi, G.H., Bailey, C., Hromockyj, A. and Kunsch, C.A.  
Enterococcus faecalis polynucleotides encoding EF059  
Patent: US 6913907-A 219 05-JUN-2005;  
Human Genome Sciences, Inc.; Rockville, MD;  
MOX;  
FEATURES  
source Location/Qualifiers  
1..1687  
/organism="unknown"  
/mol\_type="genomic DNA"  
ORIGIN  
Query Match 89.7%; Score 1687; DB 2; Length 1687;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1687; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 96 AAGAGAAATGGGAGAGGCGACAGCTCGTGAATTCACAAAAGAAATGACGATTTAC 155  
Db 1 AAGAGAAATGGGAGAGGCGACAGCTCGTGAATTCACAAAAGAAATGACGATTTAC 60  
Qy 156 AGATCCGCTTAATCAAAATAGCGGGAAGAAATGAGCGATTGATTAATCAAGACT 215  
Db 61 AGATCCGCTTAATCAAAATAGCGGGAAGAAATGAGCGATTGATTAATCAAGACT 120  
Qy 216 GCGAATGTCAGCTTATGATTTATTAACGTGACGAACGAATTTTACGACCAACGAGCGC 275  
Db 121 GCGAATGTCAGCTTATGATTTATTAACGTGACGAACGAATTTTACGACCAACGAGCGC 180  
Qy 276 AGGCGCAAGCGTTGATGACGCTTAACAGCTGCTCAAGTTTATCTCTGGAAACCTGT 335  
Db 181 AGGCGCAAGCGTTGATGACGCTTAACAGCTGCTCAAGTTTATCTCTGGAAACCTGT 240  
Qy 336 TECTCAAGGAACCAACGATGCAATGGGAATGTCACCTGTCAGTTACTTAACCAAA 395  
Db 241 TECTCAAGGAACCAACGATGCAATGGGAATGTCACCTGTCAGTTACTTAACCAAA 300  
Qy 396 TGGTAAAGATGACGTATATCAATTAAGAAACCAAAAGAGGTTGATGCTGTAC 455  
Db 301 TGGTAAAGATGACGTATATCAATTAAGAAACCAAAAGAGGTTGATGCTGTAC 360  
Qy 456 GAAATATGATGCTGCGCTTCCAGCTTACGAATGATCAAGCAACAGATGTTCTATTA 515  
Db 361 GAAATATGATGCTGCGCTTCCAGCTTACGAATGATCAAGCAACAGATGTTCTATTA 420  
Qy 516 AATATGAACAGAAATAGCGGTTGATTAATTAATCTTAATAATGTTGTCACAAATGA 575  
Db 421 AATATGAACAGAAATAGCGGTTGATTAATTAATCTTAATAATGTTGTCACAAATGA 480  
Qy 576 TGGTATTTTACATGTGAAAAAAGTAGAATGCTGTAATGAAGATTAAATGCGCAGA 635  
Db 481 TGGTATTTTACATGTGAAAAAAGTAGAATGCTGTAATGAAGATTAAATGCGCAGA 540  
Qy 636 ATTGTTTATTTCTTAATACGGAAGGCTCACAGGCACTGTAATTAATTCAGAGAGTCA 695  
Db 541 ATTGTTTATTTCTTAATACGGAAGGCTCACAGGCACTGTAATTAATTCAGAGAGTCA 600  
Qy 696 AGATGATTAATATCATGACCAACGATTAAGAAACCAACCAACGCTTATTTACTGGGA 755  
Db 601 AGATGATTAATATCATGACCAACGATTAAGAAACCAACCAACGCTTATTTACTGGGA 660  
Qy 756 AAGTTATGAATTTGGCGAAATGATTTTCAAGAGCAGAGATGACGGAAGATTAAAC 815  
Db 661 AAGTTATGAATTTGGCGAAATGATTTTCAAGAGCAGAGATGACGGAAGATTAAAC 720



|    |      |                                                                 |      |
|----|------|-----------------------------------------------------------------|------|
| QY | 816  | AGTTAAAAATCTGAGGTTGGTGTGCTATATTGTAAGAAAGTAAGTAAAGTCCAAATATATGC  | 875  |
| Db | 721  | AGTTAAAAATCTGAGGTTGGTGTGCTATATTGTAAGAAAGTAAGTCCAAATATATGC       | 780  |
| QY | 876  | AGAAATTAATTGAAATTCAAACAAAAACAATTACAAATTGAAAGCAAAACATTCAAACACC   | 935  |
| Db | 781  | AGAAATTAATTGAAATTCAAACAAAAACAATTACAAATTGAAAGCAAAACATTCAAACACC   | 840  |
| QY | 936  | TGTTGAAAAAACAGTCAAAAATGATACCTCTAAAGTTGATTAACAAACACACAGCTTAGA    | 995  |
| Db | 841  | TGTTGAAAAAACAGTCAAAAATGATACCTCTAAAGTTGATTAACAAACACACAGCTTAGA    | 900  |
| QY | 996  | TGTTAAAGATGTGGCAATTGGCGAAAAATTAAATATCAAAATTTCTGTAAATATTCATT     | 1055 |
| Db | 901  | TGTTAAAGATGTGGCAATTGGCGAAAAATTAAATATCAAAATTTCTGTAAATATTCATT     | 960  |
| QY | 1056 | GGGATTTGCACAAAGAAAGGCGAGCTTAATAAGTCAATTCGAATTTAGTTGATTA         | 1115 |
| Db | 961  | GGGATTTGCACAAAGAAAGGCGAGCTTAATAAGTCAATTCGAATTTAGTTGATTA         | 1020 |
| QY | 1116 | ACATGATCAGCCTTAACTTTTGATACGTGACTTGTGAGAGATAGCTTAATGCGTTATA      | 1175 |
| Db | 1021 | ACATGATCAGCCTTAACTTTTGATACGTGACTTGTGAGAGATAGCTTAATGCGTTATA      | 1080 |
| QY | 1176 | TGATGGGGAATACAGATATTGCTCTCGAATAATTAATGATGACTGAACAGCAATGCGTT     | 1235 |
| Db | 1081 | TGATGGGGAATACAGATATTGCTCTCGAATAATTAATGATGACTGAACAGCAATGCGTT     | 1140 |
| QY | 1236 | CAGTGTCCCGCGTTAATCCAGCGTATATCTACGCTTAACACAGCGCGCACACTAAATTT     | 1295 |
| Db | 1141 | CAGTGTCCCGCGTTAATCCAGCGTATATCTACGCTTAACACAGCGCGCACACTAAATTT     | 1200 |
| QY | 1296 | CGTTACTCTTAATGCAATTTAAATGAAAAAGCAGATCTTACGAAAGCTTTAAAAATGAGGC   | 1355 |
| Db | 1201 | CGTTACTCTTAATGCAATTTAAATGAAAAAGCAGATCTTACGAAAGCTTTAAAAATGAGGC   | 1260 |
| QY | 1356 | GAATGTTATTAACGGTCAATACCGACGACCAACACACCAACTGTGTAAGTTGTGACAGG     | 1415 |
| Db | 1261 | GAATGTTATTAACGGTCAATACCGACGACCAACACACCAACTGTGTAAGTTGTGACAGG     | 1320 |
| QY | 1416 | TGGGAACGTTTCATTTAAAGTCGATGGCGATGTGACAGGACACAGCCTTGGCGGAGC       | 1475 |
| Db | 1321 | TGGGAACGTTTCATTTAAAGTCGATGGCGATGTGACAGGACACAGCCTTGGCGGAGC       | 1380 |
| QY | 1476 | TTCCCTTGTGTCGTCGTATCAAAAACAGCGACACACAGCAATTAATTTGAAAAATCGATGAAC | 1535 |
| Db | 1381 | TTCCCTTGTGTCGTCGTATCAAAAACAGCGACACACAGCAATTAATTTGAAAAATCGATGAAC | 1440 |
| QY | 1536 | AACGAAACGCAACTTGGGTGAAAAACAAAGCTGAAGCACTATTTTACACACAGGC         | 1595 |
| Db | 1441 | AACGAAACGCAACTTGGGTGAAAAACAAAGCTGAAGCACTATTTTACACACAGGC         | 1500 |
| QY | 1596 | TGATGATTAAGTTGATATCACAGGCGCTTAAATACGTAACCTATTTTAGAAGAAACGT      | 1655 |
| Db | 1501 | TGATGATTAAGTTGATATCACAGGCGCTTAAATACGTAACCTATTTTAGAAGAAACGT      | 1560 |
| QY | 1656 | AGCTTCCTGATGATTAATGTCCTGTTAACAAATCGATTAATTTTGTCGCAATGAACATC     | 1715 |
| Db | 1561 | AGCTTCCTGATGATTAATGTCCTGTTAACAAATCGATTAATTTTGTCGCAATGAACATC     | 1620 |
| QY | 1716 | ATATGGCAACACAGAAACCTATGTTTACACAGAAAGATACCAACCAACCAAGATAC        | 1775 |
| Db | 1621 | ATATGGCAACACAGAAACCTATGTTTACACAGAAAGATACCAACCAACCAAGATAC        | 1680 |
| QY | 1776 | CTTACCT 1782                                                    |      |
| Db | 1681 | CTTACCT 1687                                                    |      |

|           |                                                                                                                          |
|-----------|--------------------------------------------------------------------------------------------------------------------------|
| ACCESSION | CQ969772                                                                                                                 |
| VERSION   | CQ969772.1                                                                                                               |
| KEYWORDS  | GI:56743733                                                                                                              |
| SOURCE    |                                                                                                                          |
| ORGANISM  | Enterococcus faecium<br>Enterococcus faecium<br>Bacteria; Firmicutes; Lactobacillales; Enterococcaceae;<br>Enterococcus. |
| REFERENCE | 1                                                                                                                        |
| AUTHORS   | Meinke, A., Nagy, E., Hanner, M. and Gelbmann, D.                                                                        |
| TITLE     | Enterococcus antigens                                                                                                    |
| JOURNAL   | Patent: WO 2004106367-A 347 09-DEC-2004;<br>Intercell AG (AT)                                                            |
| FEATURES  | location/Qualifiers                                                                                                      |
| source    | 1..1878<br>/organism="Enterococcus faecium"<br>/mol_type="unassigned DNA"<br>/db_xref="taxon:1352"                       |

ORIGIN

|                            |        |                     |           |              |
|----------------------------|--------|---------------------|-----------|--------------|
| Query Match                | 53.9%  | Score 1014.6;       | DB 2;     | Length 1878; |
| Best Local Similarity      | 71.9%; | Pred. No. 2.1e-290; |           |              |
| Matches 1342; Conservative | 0;     | Mismatches 519;     | Indels 6; | Gaps 1.      |

|    |     |                                                                  |     |
|----|-----|------------------------------------------------------------------|-----|
| QY | 1   | ATGAAAGCAATTAAAAAAGTTTGGTACACGGTTGTACTCTGTATCTAAATTTTGGCAGT      | 60  |
|    |     |                                                                  |     |
| Db | 1   | ATGAAAAATCATAAAAAAATATAACGTTATGTTAGAGATCTTTTCCTTATTTTACCATTA     | 60  |
|    |     |                                                                  |     |
| QY | 61  | TTCAACATGTATATTAGGACAAACACTGCATTTTGCAGAAAGAAATGGGAGACGCACAG      | 120 |
|    |     |                                                                  |     |
| Db | 61  | CTCAACAAACAGCTTGGCGCAAAAAAAGATGTTTGCAGAGAGACAG-----CAGCTCAA      | 114 |
|    |     |                                                                  |     |
| QY | 121 | CTCGTGAATTCACAAAAGAAAAATGACGGATTTTACCAATCCGCTTATTCGAAATATGCGGG   | 180 |
|    |     |                                                                  |     |
| Db | 115 | GTCAATCTTTCATAAAAAAGAAAAATGACTGATTTTACCCGATCTCTTATATCAAAACAGCGGG | 174 |
|    |     |                                                                  |     |
| QY | 181 | AAAGAAATAGACGAGTTTGATTAATATCAAGACCTGCAGATGTGACGTTAGATTAT         | 240 |
|    |     |                                                                  |     |
| Db | 175 | AAAGAAATAGACGAGTTTGATTAATATCAAGATTTAGCCGATATTTCATTTTCAGTTAT      | 234 |
|    |     |                                                                  |     |
| QY | 241 | AACGTGACGAACGAATTTTACGAGCAACGAGCGGCGCAGCGGCGATGACAGCTTAA         | 300 |
|    |     |                                                                  |     |
| Db | 235 | AACGTCACTCAAGAAATTTTATGCGCAACGAGATTAAGAGCGTCCGTGATGACAGCAAA      | 294 |
|    |     |                                                                  |     |
| QY | 301 | CAAGTGTCCAAAGTTTAACTTCCTGGGAAACTGTGTCTCAAGAACCAACGATGCAAT        | 360 |
|    |     |                                                                  |     |
| Db | 295 | CAAGGATGCCAGTCTTTTAACTCTCGGTACACCAAGTTGCTTACGAGAACACAGATGTGAT    | 354 |
|    |     |                                                                  |     |
| QY | 361 | GGGAATGCACATGTTCAAGTACTCTTAAAAAACAATGCTTAAAGATGACAGTGTATCCATT    | 420 |
|    |     |                                                                  |     |
| Db | 355 | GGAAATGCACATTATCTTTAAGCTTAAAAAACAATGCGAAAGATGCAGTGTCAACAGCAT     | 414 |
|    |     |                                                                  |     |
| QY | 421 | AAAGAGAACCAAAAGAGGGTGTAGTGTCTGTCTGCAATATGTGTGTGGCGTTCCCAAGT      | 480 |
|    |     |                                                                  |     |
| Db | 415 | AAAGAGAACCAAAAGAGGAGTGTCACTGCCGCAAACTAGTGTTTAGCTTTCCCTGTAT       | 474 |
|    |     |                                                                  |     |
| QY | 481 | TACGAAATGATTCACAAACAAGATGCTTCTATAATATGGAACGAAGAATTAGCGTT         | 540 |
|    |     |                                                                  |     |
| Db | 475 | TATGAGATGATTAACAACAGCAGATGTGCTTTATTAATACGGGACGAAGAACTAATACT      | 534 |
|    |     |                                                                  |     |
| QY | 541 | GTTCAATATTTATCTTAAAAATGTGTAGCCAAATGATGTAGTTTACATGTGAAAAAGTA      | 600 |
|    |     |                                                                  |     |
| Db | 535 | ATCATTCTCTAACCCCTAAAAATACAGTCGGTATGATGGAACGTTGAAGTTACAAAATTC     | 594 |
|    |     |                                                                  |     |
| QY | 601 | GGAACTGCTGAAAAATGAAGGATTTAATGTCCGCAAAATTTGTATTTCTTAAAAAGCAAAAGC  | 660 |
|    |     |                                                                  |     |
| Db | 595 | GGTACTGTCCGAAAAAGAAAGCACTAATATGAGAGAAATTTATTTATTTCTTAAAGAAAGAGA  | 654 |
|    |     |                                                                  |     |
| QY | 661 | TCACACAGGCAACGTAAATATATATCCAAAGAGTCAAAAGATGTATTTATATACATGACCAACG | 720 |
|    |     |                                                                  |     |
| Db | 655 | ACACCAAGCGTCAAAAAATACATCCAAAGTGTCAAGATGTGATTTGTACACTTTGGACMACT   | 714 |
|    |     |                                                                  |     |
| QY | 721 | GATTAAGAACCAAGCAAAAACGCTTTATTACTGGGAAAAAGTTATGAATTTGGCGAAAATGAT  | 780 |
|    |     |                                                                  |     |



Db 715 GATCAACCAAGCCAAACATTTCTATCTGCTATCTTATGATCGCAACATGAC 774  
Qy 781 TTCAAGAAAGCAGAAATGGAACGGAGAAATTAAACGTTAAAAATTTGAGTTGGTTGG 840  
Db 775 TTTGCCGAGGCACTATGAAAAAGCCAGTTGATGATTATCAATTAAAGTTGAAAA 834  
Qy 841 TATATTTTGAAGAGTAAGATCCCAATTAATGAGAAATTAATGAATAACAAAA 900  
Db 835 TATATTTTGAAGAGTAAGATCCCAATTAATGAGAAATTAATGAATAACAAATC 894  
Qy 901 ACACCAATTTACATTTGAAACCAACATCAACACCTGTTGAAAAAAACAGTCAAAAATGAT 960  
Db 895 AGCCCTTTTGAAGATCCGCAATAGCCAAACCCAGTGAAGAAAAAGCATCAAAAATGAT 954  
Qy 961 ACCTCTAAAGTTGATTAACACACACACACTTATGATGATTAAGATGTCATTTGGCGAA 1020  
Db 955 AGCTCTAAAGTTGATTAACACACACACTTATGATGATTAAGATGTCATTTGGCGAA 1014  
Qy 1021 AAAATTTAATCAAAATTTCTGTAATATTTCAATTTGGGATTCAGACAAAGAGCGAC 1080  
Db 1015 AAAATTTCAATTAAGATTTCTGCTATATCCCATTTGATGATGATTAAGAGAGAG 1074  
Qy 1081 GCTAATTAATACGTCAATTTCAATTTAGTTGATTAACATGATGACACCTTTAATCTTTGAT 1140  
Db 1075 CAACCAAGTACACAAACATTTCAATTTAGTATCTGATGACGCTGCTTTAATCAATTTGAT 1134  
Qy 1141 AACGTGACTTCTGAGAGATGATGCTTATGCTTATATGATGGGATTCAGTATGCTCT 1200  
Db 1135 AATGATTTCTTCAAGAAAGTATGCTTATATGATGATTAAGATTAAGATTCAGCCCA 1194  
Qy 1201 GAAATTTATCAAGTGTGATGACAAAGCAAAATGCTGCTGCTGCTTATCAAGGAT 1260  
Db 1195 GTAAATTTATTTCTGCTCACTGACAAACAGAGATTCAGGTTTCACTTATGATTTAT 1254  
Qy 1261 ATTCTTACGCTTAACCAAGCGCGCACTAATAATTCGTTTACTTTATGATTTAATGA 1320  
Db 1255 ATTCCTTATTAATCTCTGCGCGTACATTGAATTCGTTTATGATTTAAGAGAA 1314  
Qy 1321 AAAGCAATCTTACGAAAGGCTTTAAAAATGAGCGAATGTTGATTAAGCTTATCCGAC 1380  
Db 1315 AAAGCAATCTTACGAAAGGCTTTAAAAATGAGCGAATGTTGATTAAGCTTATCCGAC 1374  
Qy 1381 GACCAAAACCAACAACTGTTGAAGTTGACAGGCGGAAAGCTTTCAATTAAGTCAAT 1440  
Db 1375 GATCAAAACCAACCGTCACTGATGATGCTTATCGGGGCAACGATTTGTTAAAGTAT 1434  
Qy 1441 GCGCATGTCACAGCAGCAACAGCCTTGGCGGAGCTTCTTGTCTGCTGATCAAAAAC 1500  
Db 1435 GGTGACGTTATCATGACCAAAACCTGCTGAGAGAGATTCGTTGATCAAGAT 1494  
Qy 1501 AGCGACACGCAAAATTTATGAAATTCATGAAACAGAAAGCAGCACTTGGGTGAAA 1560  
Db 1495 AGTGACACGCAAAATTTATTCGATCGACCCATCCAAAGCCCTGCTGCGGATGATG 1554  
Qy 1561 ACAAAGCGTGAAGCACTACTTTTCAAGAACGCGCTGATGATTTGTTATATACAGG 1620  
Db 1555 GCGAAAGATTCAGCAACGCTTTTATCAACCAAGTAAAGCTTATGATGATGACAGAT 1614  
Qy 1621 CTTAAATACGTAACCTATTTATTAAGAAAGCTGATGCTCTGATGATTTATGCTTTGTA 1680  
Db 1615 CTTAAATACGTAACCTATTTATTAAGAAAGCTGATGCTCTGATGATTTATGCTTTGTA 1674  
Qy 1681 ACAAATCGATGATTTGTTGCTGATGAAACATCATATGCGACACAGAAACCTAGTT 1740  
Db 1675 ACAAACCGGTGATGATTTATTCATGATGAAACATCTTATGTAACAGCAGACAGTGTGAT 1734  
Qy 1741 TCACCAAGAAAGTACCAACAAACCAAGAGCTTACCTTCAACAGATGCGCAAGGA 1800  
Db 1735 TCTCTGAAAGAAATACCAATTAACCAAGAGTACCTTCTTCAACAGATGCGGATGAGGA 1794  
Qy 1801 ATCTAGTTTACTTAAAGAGGCGCGAGTCTTGTCTAATTTAGCAGAGTCTAATTTGCT 1860  
Db 1795 ATCTAGTTTACTTAAAGAGGCGCGAGTCTTGTCTAATTTAGCAGAGTCTAATTTGCT 1854

Qy 1861 AGACGTA 1867  
Db 1855 AGACGCA 1861

RESULT 12  
AR348987  
LOCUS AR348987 1995 bp DNA linear PAT 17-AUG-2003  
DEFINITION Sequence 3598 from patent US 6583275.  
ACCESSION AR348987  
VERSION AR348987.1 GI:33746032  
KEYWORDS  
SOURCE  
ORGANISM  
Unknown.  
REFERENCE  
1 (bases 1 to 1995)  
Douceite-Stamm, L.A. and Bush, D.  
Nucleic acid sequences and expression system relating to  
Enterococcus faecium for diagnostics and therapeutics  
Patent: US 6583275-A 3598 24-JUN-2003;  
Genome Therapeutics Corporation; Waltham, MA  
LOCATION/Qualifiers  
1..1995  
/organism="unknown"  
/mol\_type="genomic DNA"

ORIGIN  
Query Match 53.9%; Score 1014.6; DB 2; Length 1995;  
Best Local Similarity 71.9%; Pred. No. 2.1e-290;  
Matches 1342; Conservative 0; Mismatches 519; Indels 6; Gaps 1;

Qy 1 ATGAAGCAATTAATAAAGTTTGGTACACCGCTTATGATCTTGTATTAATTTTGCACCTT 60  
Db 118 ATGAAGCAATTAATAAAGTTTGGTACACCGCTTATGATCTTGTATTAATTTTGCACCTT 177  
Qy 61 TTCACAAGTATTAAGGACCAACATGCTGATTTGAGAGAAATGAGGAGAGCGACAG 120  
Db 178 CTCACAAAGCTTGGCGCAAAAAGTGTGTCAGAGAGACAG-----CAGCTCAA 231  
Qy 121 CTGCTGATTCACAAAAGAAATGACGATTTTACAGATCCGCTTATTCAAAATAGCGG 180  
Db 232 GTCATCCTTCATTAATAAAGAAATGACTGATTTTACCGATCTTATTCAAAACAGCGG 291  
Qy 181 AAAAGAAATGAGGAGTGTATTAATTCAGAGATGCGAGATGTCAGTTAGTATTTAT 240  
Db 292 AAAAGAAATGAGGAGTGTATTCATCAATACCAAGATTTAGCGATATTTTCAATTTCA 351  
Qy 241 AACGTGACGAAGATTTTATGAGCAACGAGCGGCGCAAGCGTGTATGATGACGTA 300  
Db 352 AACGTGACGAAGATTTTATGCGCAACGAGATTAAGAGCGTCCGTGATGACGTA 411  
Qy 301 CAAGCTGTCGAAGTTTATCTCTGGGAAACCTGTTGCTCAAGAAACACCGATGCAAT 360  
Db 412 CAAGCAGTCAGATCTTTGATCTCTGCTGATCACAGTTGCTTCAAGAAACAGATGCTGAT 471  
Qy 361 GGAATGTCACGCTTCACTTACCTTAAAAAACAATGTTAAAGTTCAGTATTAACATT 420  
Db 472 GGAATGTCACCTTATCTTAACTTAAAAAACAATGTTAAAGTTCAGTATTAACATT 531  
Qy 421 AAAGAAAGCAAAAGAGGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480  
Db 532 AAAGAAAGCAAAAGAGGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 591  
Qy 481 TAAGAAATGATCAAGCAACAGATGCTTCTTAAATATGAAACAGAAATTAAGCGGTT 540  
Db 592 TATGATGATCAACCAAGCAGATGCTCTTAAATACGGGACAGAAATTAAGCTAGTACT 651  
Qy 541 GTTCAATTTATCTTAAATATGTTGATGATGATGATGATGATGATGATGATGATGAT 600  
Db 652 ATCCATCTCTACCTTAAATAACATGCTGATGATGATGATGATGATGATGATGATGAT 711  
Qy 601 GGAATGCTGAAGAAAGATTAATGAGCGCAAGATTTGTTATTTCTAAAGCGAAGCG 660

|    |      |                                                                  |      |
|----|------|------------------------------------------------------------------|------|
| Db | 712  | GGTACTGCGGAAAAAGAAAGCACTAATATGAGAGAAATTTATTTATTTCTTAAGAAAGAAAGA  | 771  |
| Oy | 661  | TCACCAAGGCAACGTTAAATATATTCAGAGAGTCAAAAGTGGATTTATATATACATGCAACG   | 720  |
| Db | 772  | ACACCAACCGTCAAAAAATATCAATCCAAAGTGTCAACAGATGATGTCTACATCTTGCAACAT  | 831  |
| Oy | 721  | GATAAAGAAACAGCAAAACGCTTTATTTCTGGGAAAAGCTTTAGAAATTTGGGAAAAATGAT   | 780  |
| Db | 832  | GATCAACCAAAAGCCAAACATTTCTATTTCTGTCTATTTCTATGACATCGGCACACATATAC   | 891  |
| Oy | 781  | TTTCACAGAACGAGATGAAACGGAGAAATTTACAGTTAAAAATCTTGAGTTGGTTTC        | 840  |
| Db | 892  | TTTTCGCGAGCATCTATTGAAAAAGGCGAGTTGATCGTTAAATCATTTTGAAGTTGGAAAA    | 951  |
| Oy | 841  | TATATTTTGAAGAAATTAAGCTCCAAATATATGAGAAATTAATTTGAATAATCAACAA       | 900  |
| Db | 952  | TATATTTTGAAGAAATTAAGCTCTGTATTAATGCGAAAAATGATTTGAAGAAAGCAATTC     | 1011 |
| Oy | 901  | ACACCATTTTCAATTTGAAGCAAAACATCAACACCTGTGTGAAAAAAACAGTCAAAATGAT    | 960  |
| Db | 1012 | ACGCTTTTGGAGTCTGTGCAAAATGACCAACACAGTATGAAAAAGCAATCAAAAATGAT      | 1071 |
| Oy | 961  | ACCTTTAAAGTTGATTAACACACCAACGCTTAATGATGTTAAAGATGTGCAATTTGGCA      | 1022 |
| Db | 1072 | ACGCTTAAAGTTGATTAACCAACACCTCAATTTGAATGAAAAAGATGTGCAATCGTSTA      | 1131 |
| Oy | 1021 | AAAAATTAATATCAAAATTTCTGTAAATATTTCCATGTGGGGAATTCGACAAAGAAAGGCAC   | 1080 |
| Db | 1132 | AAAAATTCATATGAGATTTCTGTCAATATCCATTTAGTATTCGCTGATTAAGAAAGAAAC     | 1191 |
| Oy | 1081 | GCTAATTAATATCGTCAAAATTCATTTAGTTGATTAACATGATGACGCTTAACTTTTAT      | 1140 |
| Db | 1192 | CAAAACAAAGTACACAAACATTCAACTTATCGATTCATGTACGCTGCTTTTAACTTTGAT     | 1251 |
| Oy | 1141 | AACGGACTTCTGAGAGATATGCTTATACGTTATATGATGAGGGAATACGATATGCTCT       | 1200 |
| Db | 1252 | AATGATTTCTTCAAGAACGATGATGCTTATGCTTATATGATGAAATATGAAATTCACCCA     | 1311 |
| Oy | 1201 | GAAAAATATCAAGTACTGAACAAAGCAATGTGCTTCACTGTGCGCGTTATCCAGCGTAT      | 1260 |
| Db | 1312 | GTAATATTTCTGTCACTGAGCAACACAGCGATTCACGGTTTCAGTTGATCCGAATAT        | 1371 |
| Oy | 1261 | ATTCCTAGCTTAACCAACGCGGCGCACATTAATTTCTTTACTTTATATGCAATTTAATGA     | 1320 |
| Db | 1372 | ATTCCTTATTAATCTCTGCGGCGATTAATTAATTCGTTACTATATATCATTGTGAACAA      | 1431 |
| Oy | 1321 | AAAGCAGATCTCAACGAAAGGCTTTTAAATGATGAGCGAATGTGTAACGCTATCCGAC       | 1380 |
| Db | 1432 | AAAGAGATCCAAACCAAGAAATTTCTTAACCAAGCAATATGTCGATTAACGGCGATACAAAT   | 1491 |
| Oy | 1381 | GACCAACACACCAACCTGTGAAATGTATGACAGGTGGGAAACGTTTCATTTAAAGTCAT      | 1440 |
| Db | 1492 | GATCAAAACACACCGTCAAGTCAGTGTCTGTACTGTGGGGCAAAACGATTTGTTTAAAGTAAAT | 1551 |
| Oy | 1441 | GAGCATGTGACAGCGACACAAGCGCTTGGCGGAGCTTCTTTGTGTGTCGCGTATCAAAAC     | 1500 |
| Db | 1552 | GGTACGTTACATGACACCAACCTTGTGTGAGCAAAATTCGTGTTCTGTGATCAAGAT        | 1611 |
| Oy | 1501 | AGCGACACAGCAAAATTTTGAATATGATGAACAAACGAAAGCAGCAACTTGGGTGAA        | 1560 |
| Db | 1612 | AGTGACACAGGAAATTTATTTATGATGACCAATCCACAAAAGCGGTCAAGTGGGATATCG     | 1671 |
| Oy | 1561 | ACAAAGCTGAAGCAACTTTTATCAACAAACGCTGATGATTTAGTTGATATATCACAGGG      | 1620 |
| Db | 1672 | GCGAAAGATCCGCAACGGTTTTTACACACCAAGATTAACGGTTTATGAGATGTACAGGT      | 1731 |
| Oy | 1621 | CTTAATATCGGTACTATTTATTTAGAAAGAACTGTAGCTCTGTAGATATATGTCTTGTATA    | 1680 |
| Db | 1732 | CTAAATATATGGAACGTACTATCTGTGAAGAAACGAAAGGCCAGAAAAATATATGTTCCATTTA | 1791 |
| Oy | 1681 | ACAAATCGGATTTGAATTTGTGTCTATGACATATATATGCGACACAGAAAACTTATGTT      | 1740 |

| QY                                                               | Db   | 1792                                                         | ACAAACGGTGAAGATTACTATCGATGAAACAACTTATGTAACAGCAGCAGCTGTGATT   | 1851 |
|------------------------------------------------------------------|------|--------------------------------------------------------------|--------------------------------------------------------------|------|
| QY                                                               | Db   | 1741                                                         | TCACGAAAAAAGTACCAACAAACAAAGAGTACTTACTTCAACAGGTGGCAAGGA       | 1800 |
| QY                                                               | Db   | 1852                                                         | TCTCTGAAAAAATPACCAATPAAACAAAGGTACACTTCTTCAACAGGGGGAAGGA      | 1911 |
| QY                                                               | Db   | 1801                                                         | ATCTACGTTTACTTAGAAGTGGCGAGTCTTGTACTATTATGAGAGTCTACTTGGCT     | 1868 |
| QY                                                               | Db   | 1912                                                         | ATCTATGTGTATATTCGGTGCAGAGGTAGTCTTCTTACTGATTGGTCGACTTACTTGGCT | 1971 |
| QY                                                               | Db   | 1861                                                         | AGAGCTA 1867                                                 |      |
| QY                                                               | Db   | 1972                                                         | AGACGCA 1978                                                 |      |
| RESULT 13                                                        |      |                                                              |                                                              |      |
| AX476900                                                         |      |                                                              |                                                              |      |
| LOCUS                                                            |      |                                                              |                                                              |      |
| DEFINITION                                                       |      |                                                              |                                                              |      |
| Sequence 28 from Patent WO0212294.                               |      |                                                              |                                                              |      |
| AX476900                                                         |      |                                                              |                                                              |      |
| ACCESSION                                                        |      |                                                              |                                                              |      |
| AX476900.1                                                       |      |                                                              |                                                              |      |
| VERSION                                                          |      |                                                              |                                                              |      |
| GI:22216159                                                      |      |                                                              |                                                              |      |
| KEYWORDS                                                         |      |                                                              |                                                              |      |
| SOURCE                                                           |      |                                                              |                                                              |      |
| ORGANISM                                                         |      |                                                              |                                                              |      |
| Enterococcus faecalis                                            |      |                                                              |                                                              |      |
| Enterococcus faecalis                                            |      |                                                              |                                                              |      |
| Bacteria; Firmicutes; Lactobacillales; Enterococcaceae;          |      |                                                              |                                                              |      |
| Enterococcus.                                                    |      |                                                              |                                                              |      |
| REFERENCE                                                        |      |                                                              |                                                              |      |
| 1                                                                |      |                                                              |                                                              |      |
| ADDISON, E. and BOHNSACK, J.                                     |      |                                                              |                                                              |      |
| Group b streptococcus polypeptides nucleic acids and therapeutic |      |                                                              |                                                              |      |
| compositions and vaccines thereof                                |      |                                                              |                                                              |      |
| Patent: WO 0212294-A 28 14-FEB-2002.                             |      |                                                              |                                                              |      |
| ST. JUDE CHILDREN'S RESEARCH HOSPITAL (US) ; University of Utah  |      |                                                              |                                                              |      |
| Research Foundation (US)                                         |      |                                                              |                                                              |      |
| FEATURES                                                         |      |                                                              |                                                              |      |
| source                                                           |      |                                                              |                                                              |      |
| 1. .2199                                                         |      |                                                              |                                                              |      |
| Location/Qualifiers                                              |      |                                                              |                                                              |      |
| /organism="Enterococcus faecalis"                                |      |                                                              |                                                              |      |
| /mol_type="unassigned DNA"                                       |      |                                                              |                                                              |      |
| /db_xref="taxon:1351"                                            |      |                                                              |                                                              |      |
| ORIGIN                                                           |      |                                                              |                                                              |      |
| Query Match                                                      |      |                                                              |                                                              |      |
| Best Local Similarity                                            |      |                                                              |                                                              |      |
| 31.6%; Score 595; DB 2; Length 2199;                             |      |                                                              |                                                              |      |
| 100.0%; Pred. No. 1.9e-165;                                      |      |                                                              |                                                              |      |
| Matches 595; Conservative 0; Mismatches 0; Indels 0; Gaps 0;     |      |                                                              |                                                              |      |
| QY                                                               | 1287 | ACTTAATTCGTTTACTTATGCACTTTAAATGAAAAAGCAGATCTCTAGAAAGGCTTTAA  | 1346                                                         |      |
| Db                                                               | 1    | ACTTAATTCGTTTACTTATGCACTTTAAATGAAAAAGCAGATCTCTAGAAAGGCTTTAA  | 60                                                           |      |
| QY                                                               | 1347 | AAATGAGCGAATGTTGATTAACGGTCATPACCGAAGCAACCAACCACTGTGTAAGT     | 1406                                                         |      |
| Db                                                               | 61   | AAATGAGCGAATGTTGATTAACGGTCATPACCGAAGCAACCAACCACTGTGTAAGT     | 120                                                          |      |
| QY                                                               | 1407 | TGTGACAGGTGGGAAACGTTTCATTTAAAGTCGATGGCGATGTGAACGACACAGCCTT   | 1466                                                         |      |
| Db                                                               | 121  | TGTGACAGGTGGGAAACGTTTCATTTAAAGTCGATGGCGATGTGAACGACACAGCCTT   | 180                                                          |      |
| QY                                                               | 1467 | GGCGGAGGCTTCCTTGTGCGTGGTGAATCAAAACAGGACACACAAATTTATTTGAAAT   | 1526                                                         |      |
| Db                                                               | 181  | GGCGGAGGCTTCCTTGTGCGTGGTGAATCAAAACAGGACACACAAATTTATTTGAAAT   | 240                                                          |      |
| QY                                                               | 1527 | CGATGAAACGAAGAGCAACCTTGGGTGAAAAACAAAGCTGAGCAACTACTCTTTAC     | 1586                                                         |      |
| Db                                                               | 241  | CGATGAAACGAAGAGCAACCTTGGGTGAAAAACAAAGCTGAGCAACTACTCTTTAC     | 300                                                          |      |
| QY                                                               | 1587 | AACCAAGGCTGATGATTAAGTTGATATCACAGGCTTTAAATACGATACCTATTATTAGA  | 1646                                                         |      |
| Db                                                               | 301  | AACCAAGGCTGATGATTAAGTTGATATCACAGGCTTTAAATACGATACCTATTATTAGA  | 360                                                          |      |
| QY                                                               | 1647 | AGAAATGTAGCTCCTGATGATTAATGCTTGTGTTAACAAAATCGATTGAATTTGTGTCAA | 1706                                                         |      |
| Db                                                               | 361  | AGAAATGTAGCTCCTGATGATTAATGCTTGTGTTAACAAAATCGATTGAATTTGTGTCAA | 420                                                          |      |
| QY                                                               | 1707 | TGAACATCATATGCGACACAGAAAAACCTAGTTTACACGAAAAAAGTACCAACCAACA   | 1766                                                         |      |

Db 421 TGAACATCATATGACACACAGAAAGCTAGTTCCAGAAAGAGTCCAAACAAACAA 480

Qy 1767 CAAGGATACCTTACCTTCAACAGTGGCAGAGAAAGTACGTTACTTGAAGAGTGGCC 1826

Db 481 CAAAGGATACCTTACCTTCAACAGTGGCAGAGAAAGTACGTTACTTGAAGAGTGGCC 540

Qy 1827 AGCTTCTGCTACTTATTGTCAGAGTCTACTTGTCTAGACGTAGAAAAGAAATGCT 1861

Db 541 AGCTTCTGCTACTTATTGTCAGAGTCTACTTGTCTAGACGTAGAAAAGAAATGCT 595

RESULT 14  
AC181263/c 150038 bp DNA linear HTG 01-FEB-2006

LOCUS Strongylocentrotus purpuratus clone R3-4012A17, WORKING DRAFT

DEFINITION AC181263

ACCESSION AC181263.1 GI:86263499

VERSION HTG; HTGS\_PHASE1; HTGS\_DRAFT; HTGS\_POOLED\_CLONE; HTGS\_ENRICHED.

KEYWORDS Strongylocentrotus purpuratus

SOURCE Strongylocentrotus purpuratus

ORGANISM Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa; Echinoidea; Euechinoidea; Echinoidea; Strongylocentrotus; Strongylocentrotidae; Strongylocentrotus.

REFERENCE 1 (bases 1 to 150038)

AUTHORS Shen, Y., Abraham, K.K., Abulimiti, A., Adams, C.Q., Aduba, G., Allen, C.C., Alsbrooks, S.L., Anosike, U.S., Archer, P.M., Barron, S.R., Bell, A.V., Bell, S.N., Bernaducci, A.R., Bickham, C., Biewald, K., Blyth, P.R., Bunay, C.J., Canada, A., Cardenas, V., Carter, K., Chacko, J., Chandrasekhar, M.N., Chavez, A., Chavez, D., Chen, G., Chen, R., Chu, H., Clerc, Blankenburg, K.P., Cockrell, R., Cooper, J., Coyle, M.D., Cree, A., Cueto, C.B., Curry, S.M., Dai, W., Dao, M.D., Davila, M., Davis, C., Davy-Carroll, L., Del Fierro, P., Demen, R., Denson, S., Ding, Y., Dinh, H.H., Donlin, J.E., Dugan-Rocha, S., Dunn, A.M., Durbin, K.J., Eboong, V.E., Egan, A., Espinosa, V.C., Fa, M., Fernandez, S., Fernando, P.R., Ferrer, A.R., Flagg, N., Forbes, L.D., Fowler, R.G., Fu, Q., Fuh, B., Gabisi, R.A., Gaidarhanan, M., Garner, J., Garcia, I.I., R.M., Garcia, A.M., Gonzalez, S.M., Garner, T.T., Ghose, S., Gingles, M., Gonzalez-Garay, M.L., Guevara, W.V., Haaland, W.C., Heberlein, K.A., Hagena, B.J., Hall, O., Hamid, H., Hamilton, K.A., Hampton, O.A., Hargreaves, B.A., Harris, R.A., Havlak, P., Hawes, A.C., Hawkins, E.S., Haynes, S.J., Hemphill, L., Hernandez, J., Hines, S., Hirani, K., Hitchens, M.E., Hodgson, A.V., Hognes, M.E., Holder, M., Hollins, B., Howell, L.L., Hulik, S.W., Hume, J., Jackson, A., Jackson, L.R., Jacob, S.K., Jhangiani, S.N., Jiang, H., Johnson, B., Johnson, R., Joseph, V., Joy, C., Kalket, P.B., Kalafatis, K.J., Kalu, J.B., Kang, Y., Keebler, J., Khan, Z.M., Kidwai, S., King, L.M., Kisamo, H., Kovar, C.L., Kowals, A.N., Kowis, C.R., Lago, L.A., Lago, M.T., Lai, C., Lata, F., Le, T.T., Lee, S.L., Lee, T.W., Legall, I.I., F.H., Lemon, S.J., Lewis, L.R., Li, B., Li, Y., Li, Z., Linnell, M.A., Liu, J., Liu, W., Liu, Y., Liu, Y., Llyanage, D., London, P., Lopez, J., Lorenshue, L.M., Lozano, R.J., Luc, T., Madu, R.C., Maheshwari, M., Maheshwari, R., Malloy, K., Mansouri, D.L., Martinez, E., Matejkova, P., Mathew, T., McCaulley, S.K., McPherson, J.D., Mercado, C., Mercado, I.C., Metcalf, M.L., Millin, A., Mlasevle, A., Morgan, M.B., Morris, S., Mundasa, M., Murray, D.D., Muzny, D.M., Nazareth, L.V., Ngo, D.N., Nguyen, H.T., Nguyen, N.B., Nguyen, P.O., Nwaokeme, O.O., Obregon, M., Odeh, E.A., Okonko, F., Okunonu, G.O., Okunonu, K.C., Onyenekwe, J., Parish, B.J., Parker, D.N., Parra, A.A., Patel, N.K., Patel, B.M., Patel, R.R., Paul, H.A., Perez, A., Perez, L.M., Perez, Y.Y., Pham, T.L., Player, E.J., Primus, E.L., Pu, L., Puzo, M., Purkiss, C., Qin, X., Quiroz, J.B., Rabata, D., Rachlin, E.K., Ren, Y., Richards, S., Rojas, A., Ruiz, S., Sabo, A., Santibanez, J., Savory, G.G., Scherer, S.R., Schneider, B.W., Sebasingiri, R., Sexon, M.M., Shen, H., Sleson, I., Sneed, A.J., Sodergren, E., Song, X., Sorrell, R.P., Svatek, A.E., Taylor, B.W., Taylor, T.R., Thelner, R., Thomas, N., Thorm, R.D., Thornton, R.D., Tong, M.Y., Trejos, Z.Y., Umanai, K., Vargio, C.B., Valtathil, S., Vega, R.A., Villaseana, D., Volkov, A., Walker, D.L., Wang, Q., Wang, S., Warren, J.T., Watt, J.E., Wei, X., Wheeler, D.A., White, C.S., Williams, Jr, R.L., Williams, A.C.,

# TITLE JOURNAL REFERENCE AUTHORS TITLE JOURNAL

## COMMENT

Williams, G.A., Williams, J.D., Wilson, K., Woodworth, J.R., Worley, K.C., Wright, R.A., Wu, J., Wu, W., Yakub, S., Yerrapragada, S., Yu, F., Yuan, D.T., Yuan, Y., Zhang, Y., Zhang, Z., Zhou, J., Zhu, Y., Weinstock, G. and Gibbs, R.A.

Unpublished  
Direct Submission  
2 (bases 1 to 150038)

Worley, K.C.  
Direct Submission  
Submitted (01-FEB-2006) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

Genome Center  
Center: Baylor College of Medicine  
Center code: BCM  
Web site: <http://www.hgsc.bcm.tmc.edu/>  
Contact: hgsc-help@bcm.tmc.edu  
Project Information  
Center clone name: R3-4012A17  
Sequencing Vector: pUC18  
Cloning Vector: pBACe3.6  
Chemistry: Dye-terminator Big Dye  
Summary Statistics  
Estimated insert size: 1 sum-of-contigs estimation  
Quality coverage: x in Q20 bases; sum-of-contigs estimation

NOTES: The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rac/>). The BAC reads were identified through a deconvolution of an array containing pools of BACs. Due to the incorporated WGS reads, there may be contigs that consist entirely of whole genome shotgun sequence reads and the sequence may extend beyond the ends of the clone. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the features table.

NOTE: This is a 'working draft' sequence. It currently consists of 13 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1 21261: contig of 21261 bp in length

21262 21271: gap of 10 bp

21272 31480: contig of 10209 bp in length

31481 31490: gap of 10 bp

31491 45811: contig of 14321 bp in length

45812 45821: gap of 10 bp

45822 62070: contig of 16249 bp in length

62071 62080: gap of 10 bp

62081 70393: contig of 8313 bp in length

70394 70493: gap of unknown length

70494 84032: contig of 13339 bp in length

84033 84042: gap of 10 bp

84043 106287: contig of 22245 bp in length

106288 106297: gap of 10 bp

106298 124992: contig of 18695 bp in length

124993 125002: gap of 10 bp

125003 130651: contig of 5649 bp in length

130652 130751: gap of unknown length

130752 138577: contig of 7826 bp in length

138578 138587: gap of 10 bp

138588 139792: contig of 1205 bp in length

139793 139892: gap of unknown length

139893 144331: contig of 4339 bp in length

144332 144331: gap of unknown length

144332 150038: contig of 5707 bp in length.

Location/Qualifiers

1. 150038

/organism="Strongylocentrotus purpuratus"

/mol\_type="genomic DNA"

/db\_xref="taxon:7668"



RESULT 15  
AC1178783/C  
LOCUS  
DEFINITION  
ACCESSION  
AC1178783  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS

AC1178783 204539 bp DNA linear HTG 31-JAN-2006  
Strongylocentrotus purpuratus clone R3-21G01, WORKING DRAFT  
SEQUENCE, 31 unordered pieces.  
AC1178783  
HTG, HTGS\_PHASE1, HTGS\_DRAFT, HTGS\_POOLED\_CLONE, HTGS\_ENRICHED.  
Strongylocentrotus purpuratus  
Strongylocentrotus purpuratus  
Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;  
Echinoidae; Euechinoidae; Echinacea; Echinoida;  
Strongylocentrotidae; Strongylocentrotus.  
1 (bases 1 to 204539)  
Shen, Y., Abraham, K.K., Abulimiti, A., Adams, C.Q., Aduba, G.,  
Allen, C.C., Alstrooks, S.L., Anosike, U.S., Archer, P.M.,  
Airdondo, H.H., Altway, T., Bandaranaike, D.P., Bangura, L.,  
Barton, S.R., Bell, A.V., Bell, S.N., Beraducci, A.R., Bicham, C.,  
Biswal, K., Blyth, P.R., Buhay, C.J., Canada, A., Cardenas, V.,  
Carter, K., Chacko, J., Chandrasekhar, M.N., Chavez, A., Chavez, D.,  
Chen, G., Chen, R., Chu, H., Clerc Blankenburg, K.P., Cockrell, R.,  
Cooper, J.A., Coyle, M.D., Cree, A., Cuello, C.B., Curry, S.M., Dai, W.,  
Dao, M.D., Davila, M., Davis, C., Davy-Carroll, L., Del Fierro, P.,  
Demen, R., Denson, S., Ding, Y., Dinu, H.H., Donlin, J.E.,  
Dugan-Rocha, S., Dunn, A.M., Durbin, K.J., Eban, V.E., Egan, A.,  
Espinoza, V.C., Fa, M., Fernandez, S., Fernando, P.R., Ferrier, A.R.,  
Flagg, N., Forbes, L.D., Fowler, R.G., Fu, Q., Fuh, E., Gabisi, R.A.,  
Ganardhanan, M., Garner, J., Garcia, I.I., R.M., Garcia, A.M.,  
Garcia, S.M., Garner, T.T., Ghouse, S., Gindras, M.,  
Gonzalez-Garay, M.L., Guevara, W.V., Haaland, W.C., Heberlein, K.A.,  
Hagens, B.J., Hall, O., Hamid, H., Hamilton, K.A., Hampton, O.A.,  
Hartes, B.A., Harris, R.A., Havlik, P., Hawes, A.C., Hawkins, E.S.,  
Haynes, S.J., Hemphill, L., Hernandez, J., Hines, S., Hirani, K.,  
Hithens, M.E., Hodgson, A.V., Hogue, M.E., Holder, M., Hollins, B.,  
Howell, L.L., Huylk, S.W., Hume, J., Jackson, A., Jackson, L.R.,  
Jacob, S.K., Jhangiani, S.N., Jiang, H., Johnson, B., Johnson, R.,  
Joishi, V., Joy, C., Kaikal, F.B., Kalafatis, K.J., Kalu, J.B., Kang, Y.,  
Keeler, J., Khan, Z.M., Kidwai, S., King, L.M., Kisman, H., Kovac, C.L.,  
Kowals, A.N., Kowals, C.R., Lago, L.A., Lago, M.T., Lai, C., Lare, F.,  
Le, T.T., Lee, S.L., Lee, T.W., Legall, I.I., F.H., Lemon, S.J.,  
Lewis, L.R., Li, B., Li, Y., Li, Z., Linell, M.A., Liu, J., Liu, W.,  
Liu, Y., Liu, Y., Lijang, D., London, P., Lopez, J., Lorensen, L.M.,  
Lozdo, R.J., Luc, T., Madu, R.C., Maheshwari, M., Maheshwari, L.M.,  
Malloy, K., Mansouri, D.L., Martinez, E., Matejkova, P., Mathew, T.,  
McCauley, S.K., McPherson, J.D., Mercado, C., Mercado, I.C.,  
Meeker, M.L., Millin, A., Milosavljevic, A., Morgan, M.B., Morris, S.,  
Mundasa, M., Murray, D.D., Muzny, D.M., Nazareth, L.V., Ngo, D.N.,  
Nguyen, H.T., Nguyen, N.B., Nguyen, P.Q., Nwaokeme, O.O.,  
Obregon, M., Odeh, E.A., Okonko, F., Okunolu, G.O., Okunolu, K.C.,  
Oyeneke, J., Parish, B.J., Parker, D.N., Parra, A.A., Pasternak, S.,  
Patel, B.M., Patel, R.R., Paul, H.A., Perez, A., Perez, L.M.,  
Perez, Y.Y., Pham, T.L., Player, E.J., Primus, B.L., Pu, L., Puzo, M.,  
Purtilis, C., Qin, X., Quiroz, J.B., Rabata, D., Rachlin, E.K., Ren, Y.,  
Richards, S., Rojas, A., Ruiz, S., Sabo, A., Santibanez, J.,  
Savory, G.G., Scherer, S.E., Schneider, B.W., Sebasigiri, R.,  
Sexon, M.M., Shen, H., Sison, I., Sneed, A.J., Sodergren, E., Song, X.,  
Sorrell, R.P., Svatek, A.F., Taylor, E.W., Taylor, T.R., Theuber, R.,  
Thomas, N., Thoren, R.D., Thornton, R.D., Tong, M.Y., Tejos, Z.Y.,  
Umanis, K., Vargo, C.E., Valtachli, S., Vega, R.A., Villalana, D.,  
Volkov, A., Walker, D.L., Wang, Q., Wang, S., Warren, J.T., Watt, J.E.,  
Wei, X., Wheeler, D.A., White, C.S., Williams, Jr, R.L., Williams, A.C.,  
Williams, G.A., Williams, J.D., Wilson, K., Woodworth, J.R.,  
Worley, K.C., Wright, R.A., Wu, J., Wu, W., Yakub, S., Yerragada, S.,  
Yu, F., Yuan, D.T., Yuan, Y., Zhang, J., Zhang, L., Zhang, Z., Zhou, J.,  
Zhu, Y., Weinstein, G. and Gibbs, R.A.

TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL

2 (bases 1 to 204539)  
Unpublished  
Direct Submission  
Worley, K.C.  
Submitted (31-JAN-2006) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA

## COMMENT

----- Genome Center  
Center: Baylor College of Medicine  
Center code: BCM  
Web site: <http://www.hgsc.bcm.tmc.edu/>  
Contact: [hgsc-help@bcm.tmc.edu](mailto:hgsc-help@bcm.tmc.edu)  
Project Information  
Center clone name: R3-21G01  
Sequencing Vector: pUC18  
Cloning Vector: pBACE3.6  
Chemistry: Dye-terminator Big Dye  
Summary Statistics  
Estimated insert size: 1 sum-of-contigs estimation  
Quality coverage: x in Q20 bases: sum-of-contigs estimation

NOTE: The sequence in this assembly is a combination of BAC based  
reads and whole genome shotgun sequencing reads assembled using  
Atilas (<http://www.hgsc.bcm.tmc.edu/projects/atlas/>). The BAC reads  
were identified through a deconvolution of an array containing  
pools of BACs. Due to the incorporated WGS reads, there may be  
contigs that consist entirely of whole genome shotgun sequence  
reads and the sequence may extend beyond the ends of the clone.  
Both end sequences and whole genome shotgun sequence only contigs  
will be indicated in the features table.  
NOTE: This is a 'working draft' sequence. It currently  
\* consists of 31 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

1 14737: contig of 14737 bp in length  
14738: contig of 972 bp  
15710: contig of 2689 bp in length  
18399: gap of 10 bp  
18409: contig of 4981 bp in length  
23389: gap of 10 bp  
23399: contig of 248 bp in length  
23400: contig of 248 bp in length  
23647: gap of 10 bp  
23648: gap of 10 bp  
23658: contig of 10343 bp in length  
34001: contig of 5108 bp in length  
34011: gap of 10 bp  
39119: contig of 8666 bp in length  
39129: gap of 10 bp  
39129: contig of 8666 bp in length  
47795: gap of 10 bp  
47804: gap of 10 bp  
47805: contig of 1100 bp in length  
48904: gap of unknown length  
49004: gap of unknown length  
49005: contig of 2214 bp in length  
51218: gap of unknown length  
51219: contig of 12058 bp in length  
51319: gap of 10 bp  
51319: contig of 12058 bp in length  
51319: gap of 10 bp  
63376: gap of 10 bp  
63377: contig of 3967 bp in length  
63387: gap of 10 bp  
67353: contig of 9729 bp in length  
67364: gap of 10 bp  
77092: contig of 9729 bp in length  
77093: gap of 10 bp  
77102: gap of 10 bp  
77103: contig of 1541 bp in length  
92743: gap of 10 bp  
92744: gap of 10 bp  
92753: gap of 10 bp  
103572: contig of 10819 bp in length  
103573: gap of unknown length  
105208: contig of 1536 bp in length  
105218: gap of 10 bp  
105219: contig of 1544 bp in length  
106802: gap of unknown length  
106803: gap of unknown length  
106903: contig of 1909 bp in length  
108811: gap of 10 bp  
108822: gap of 10 bp  
139372: contig of 30551 bp in length  
139472: gap of unknown length  
140424: contig of 952 bp in length  
141570: gap of 1246 bp in length  
141671: gap of 1336 bp in length  
143006: contig of 1336 bp in length  
143107: gap of unknown length  
144509: contig of 1403 bp in length  
144510: gap of unknown length

|   |         |                              |
|---|---------|------------------------------|
| * | 14535:  | contig of 626 bp in length   |
| * | 145236: | gap of unknown length        |
| * | 145336: | contig of 1415 bp in length  |
| * | 146751: | gap of unknown length        |
| * | 146850: | contig of 789 bp in length   |
| * | 147639: | gap of unknown length        |
| * | 147640: | contig of 1249 bp in length  |
| * | 147740: | gap of unknown length        |
| * | 148988: | contig of 839 bp in length   |
| * | 148989: | gap of unknown length        |
| * | 149089: | contig of 839 bp in length   |
| * | 149928: | gap of unknown length        |
| * | 150027: | contig of 4631 bp in length  |
| * | 150028: | gap of unknown length        |
| * | 154658: | contig of 1836 bp in length  |
| * | 154659: | gap of unknown length        |
| * | 154759: | contig of 1836 bp in length  |
| * | 156594: | gap of unknown length        |
| * | 156595: | contig of 1142 bp in length  |
| * | 157836: | gap of unknown length        |
| * | 157935: | contig of 9763 bp in length  |
| * | 157837: | gap of unknown length        |
| * | 157937: | contig of 9763 bp in length  |
| * | 167799: | gap of unknown length        |
| * | 167800: | contig of 36740 bp in length |
| * | 204539: | contig of 36740 bp in length |

| FEATURES | Location/Qualifiers |
|----------|---------------------|
| source   | 1. .204539          |

```

/misc_feature 1. .14737 /note="assembly_name:Contig28"
/misc_feature 14738. .15709 /note="assembly_name:gap"
/misc_feature 14738. .15709 /estimated_length=972
/misc_feature 15710. .16398 /note="assembly_name:Contig16"
/misc_feature 18359. .18408 /note="assembly_name:gap"
/misc_feature 18399. .18408 /estimated_length=10
/misc_feature 18409. .23389 /note="assembly_name:Contig21"
/misc_feature 23350. .23359 /note="assembly_name:gap"
/misc_feature 23350. .23359 /estimated_length=10
/misc_feature 23400. .23647 /note="assembly_name:Contig1"
/misc_feature 23648. .23657 /note="assembly_name:gap"
/misc_feature 23648. .23657 /estimated_length=10
/misc_feature 23658. .34000 /note="assembly_name:Contig24"
/misc_feature 34001. .34010 /note="assembly_name:gap"
/misc_feature 34001. .34010 /estimated_length=10
/misc_feature 34011. .39118

```

|                       |        |              |          |            |      |        |
|-----------------------|--------|--------------|----------|------------|------|--------|
| Query Match           | 5.0%;  | Score        | 94.6;    | DB         | 12;  | Length |
| Best Local Similarity | 44.0%; | Pred. No.    | 1.4e-16; |            |      |        |
| Matches               | 455;   | Conservative | 0;       | Mismatches | 574; | Indels |

[illegible]

|    |        |                                                                |        |
|----|--------|----------------------------------------------------------------|--------|
| OY | 278    | CCGACGGCTTGATCAGCTAAACAAGCTGTCCAAGTTTAACTCCTGGGAACCTGTTG       | 337    |
| Dd | 133413 | TAACTTAAGAAGAAAGAGAGAGAGAGAGAGTAGTATTAAAGACGTAAAAGAAAGAAAGAG   | 133355 |
| OY | 338    | CTCAAGAACCAACCAGTCCAAATGGGAATGTCACCTGTTCACTTAATAAACAATAATG     | 397    |
| Dd | 133353 | AAGCGTAAAAAGAAAGAGTAGTAGAGAAAGATTAAGAAAGAAAGAGAGTATTAAAAAGAAAG | 133294 |
| OY | 398    | GTAAGAATGCAGTGTATACATTAAAGAAAGAACAAAGAGGGGTGTATGTTCTGTACGA     | 457    |
| Dd | 133293 | AAAAAGAGAAAGAGAGAGAGAGAGAGAGTAAAAAGAGAGATGTATTTAAAAAGAGTAAAAAG | 133233 |
| OY | 458    | ATATGCTGTGGCGTTCCCACTTTACGAATATGATCAAGCAACAGATGGTTCCTATTAAT    | 517    |
| Dd | 133233 | AAGAAAGAGAAAGCACTAAAAAGAAAGATGTATTAAGAAAGTAAAAAGAAAGAAAGAGAG   | 133174 |
| OY | 518    | ATGCAACAGAAAGATTAGCGCTTGTTCATATTATCTTAAAAATGTGTGCCAATGATG      | 577    |
| Dd | 133173 | AAGAAAGAGTAAAGAAAGAGTAGTATTAAGAGAGAGAGAGAGAGAGTATTAAGAGAGAGTAG | 133113 |

|    |        |                                                                  |        |
|----|--------|------------------------------------------------------------------|--------|
| QY | 578    | GTAGTTTACATGTGAAAAAAGTAGAACTGTGAATAATGAAGATTAAATGGCGCAAT         | 637    |
| Db | 133113 | TAAAGAAGATAAAGAGAGAGAAAGCAGTAAAGAGAGAGATGATTAAGAAGATAAAG         | 133054 |
| QY | 638    | TTGTATTTCCTAAAAGCCAGAGCTCACCAGGCACGTAAATATATCCAAAGAGTCAAAG       | 697    |
| Db | 133053 | AAG-----AAGAAAGAGAGAGAGAGCGGTAAAGAGAGAGAGTGTAGAGAGAGTAAAG        | 133000 |
| QY | 698    | ATGATTTATATATCATGTGACAAACGAGTAAAGAACAAAGCAAAACGCTTATTTACTGGGAAA  | 757    |
| Db | 132999 | AAGAAAGAGTTATTAAGAAAAAAGAAAAAGAAAGAGAGAGAGAGAGAGAGTAAAGAGAG      | 132940 |
| QY | 758    | GTTATGAAAATTCGCGCAAAATGATTTTCACAGAGACAGAGATGGAACGGAGAAATTAAACAG  | 817    |
| Db | 132939 | AAGTAGTAAAGAAAGTAAAGAAAGAGAGAGAGAGAGAGTAAAGAGAGAGTGTAAAGAGAG     | 132890 |
| QY | 818    | TTAAATAATCTTGAGGTTGGTTCGTATATTTTGAAGAGATAAAGGCTCCAAATATATGACAG   | 877    |
| Db | 132879 | TAAAGAGAGAGAGAGAGAGAGAGAGAGTAAAGAGAGAGAGAGAGTAAAGAGAGTAAAGAGAGAG | 132820 |
| QY | 878    | AATTATATGAAAAATCAAAACAAAACACCAATTTTCAATTGAGCAAAATCAAAACACCTG     | 937    |
| Db | 132819 | AAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGTATGAGAGAGAGTAAAGAGAGAGAGAG | 132760 |
| QY | 938    | TTGAAAACAAACAGTCAAAAAATGATACCTTAAGATTATATAAACAAACACCAAGCTAGATG   | 997    |
| Db | 132759 | TTATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGTAAAGAGAGAGAGTAAAG     | 132700 |
| QY | 998    | GTAAGATGTGGCAATTGGCGAAAAAATTAATATATCAAAATTTCTGTATATATTCATTGG     | 1057   |
| Db | 132699 | AAGAGATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGTAAAGAGAGAGAGAG     | 132640 |
| QY | 1058   | GGATTGCAGACAAAGAGGCGACGCTATATTAATATAGTCAAAATTCATTTTGTGATTAAC     | 1117   |
| Db | 132639 | AAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGTAAAGAGAGAGAGAGAGAGTAA      | 132580 |

|    |                              |
|----|------------------------------|
| Qy | 1118 ATGATGCAGCCTTAA 1132    |
|    |                              |
| Db | 132579 AAGAAGAAGTAGTAA 13265 |

Search completed: July 5, 2006, 20:23:46  
Job time : 7258 secs

GenCore version 5.1.9  
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 5, 2006, 18:22:58 ; Search time 6165 Seconds  
(without alignments) 17061.519 Million cell updates/sec

Title: US-10-661-809A-12

Perfect score: 1881  
Sequence: 1 atgaagcaatcaaaaaaagt.....gacgtagaagaagaatgct 1881

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 48236798 seqs, 27959665780 residues

Total number of hits satisfying chosen parameters: 96473596

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

EST:\*  
1: gb\_est1:\*  
2: gb\_est3:\*  
3: gb\_est4:\*  
4: gb\_est5:\*  
5: gb\_est6:\*  
6: gb\_hic:\*  
7: gb\_est2:\*  
8: gb\_est7:\*  
9: gb\_est8:\*  
10: gb\_est9:\*  
11: gb\_g881:\*  
12: gb\_g882:\*  
13: gb\_g883:\*  
14: gb\_g884:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description                 |
|------------|-------|-------------|--------|----|-----------------------------|
| C 1        | 496.4 | 26.4        | 500    | 11 | B07882 YX102T7 PBI          |
| C 2        | 492.6 | 26.2        | 500    | 11 | B07880 B07886 YX106T7 PBI   |
| C 3        | 492   | 26.2        | 500    | 11 | B07866 YX106T7 PBI          |
| C 4        | 434.8 | 23.1        | 450    | 11 | B07865 B07881 YX102T3 PBI   |
| C 5        | 430.2 | 22.9        | 450    | 11 | B07889 YX98T3 PBI           |
| C 6        | 380   | 20.2        | 400    | 11 | B07889 YX98T3 PBI           |
| C 7        | 91.2  | 4.8         | 908    | 11 | AZ548467 ENTEK30TR          |
| C 8        | 88.8  | 4.7         | 843    | 11 | AZ551618 ENTEK30TR          |
| C 9        | 79    | 4.2         | 1084   | 14 | AL049509 T7 end of          |
| C 10       | 78.2  | 4.2         | 1811   | 12 | CG753732 P048-4-G0          |
| C 11       | 78    | 4.1         | 877    | 11 | AZ531291 ENTEK30TR          |
| C 12       | 77.6  | 4.1         | 1416   | 9  | DN711652 ENTEK30TR          |
| C 13       | 77.2  | 4.1         | 912    | 11 | AZ551092 ENTEK30TR          |
| C 14       | 77    | 4.1         | 879    | 14 | CNS01JRG AL147405 Anopheles |
| C 15       | 76.4  | 4.1         | 993    | 7  | B8898130 B8898130           |
| C 16       | 76.4  | 4.1         | 1496   | 12 | CU073855 CH216-130          |
| C 17       | 76    | 4.0         | 1104   | 1  | AJ925669 AJ925669           |
| C 18       | 75.8  | 4.0         | 1981   | 12 | CU082000 CH216-165          |
| C 19       | 75.6  | 4.0         | 1255   | 14 | AG435644 Muv muscu          |

|      |      |     |      |    |          |                     |
|------|------|-----|------|----|----------|---------------------|
| C 20 | 75   | 4.0 | 1210 | 12 | CG751315 | CG751315 P045-4-B0  |
| C 21 | 75   | 4.0 | 1687 | 1  | AM090701 | AM090701 AM090701   |
| C 22 | 74.4 | 4.0 | 1300 | 2  | BM468018 | BM468018 AGENCOURT  |
| C 23 | 74.2 | 3.9 | 1053 | 3  | BUS08694 | BUS08694 AGENCOURT  |
| C 24 | 73.8 | 3.9 | 868  | 11 | BH161951 | BH161951 ENTEK30TR  |
| C 25 | 73.8 | 3.9 | 1348 | 9  | DN721897 | DN721897 CNB140-H1  |
| C 26 | 73.2 | 3.9 | 1896 | 10 | DV780693 | DV780693 Hw PAT 11  |
| C 27 | 73   | 3.9 | 1345 | 14 | AG429454 | AG429454 Muv muscu  |
| C 28 | 72.8 | 3.9 | 949  | 4  | CA465606 | CA465606 AGENCOURT  |
| C 29 | 72.8 | 3.9 | 1687 | 12 | CU082679 | CU082679 CH216-169  |
| C 30 | 72.6 | 3.9 | 1101 | 14 | CNS0039G | AL063621 Drosophila |
| C 31 | 72.6 | 3.9 | 1260 | 13 | CU491610 | CU491610 SA11 559   |
| C 32 | 72.6 | 3.9 | 1753 | 10 | DR131208 | DR131208 49217594   |
| C 33 | 72.4 | 3.8 | 905  | 11 | AZ550256 | AZ550256 ENTEK30TR  |
| C 34 | 72.4 | 3.8 | 1058 | 4  | CD048666 | CD048666 AGENCOURT  |
| C 35 | 72.2 | 3.8 | 1493 | 12 | CU078589 | CU078589 CH216-151  |
| C 36 | 72   | 3.8 | 964  | 13 | CW957121 | CW957121 TCBS2.1C   |
| C 37 | 72   | 3.8 | 1114 | 9  | DN656077 | DN656077 CEC21-A11  |
| C 38 | 72   | 3.8 | 1594 | 12 | CU038406 | CU038406 CH216-46A  |
| C 39 | 72   | 3.8 | 1843 | 14 | AG435185 | AG435185 Muv muscu  |
| C 40 | 71.8 | 3.8 | 1256 | 12 | CU119201 | CU119201 ISB1-76J1  |
| C 41 | 71.8 | 3.8 | 1459 | 12 | CU040025 | CU040025 CH216-49A  |
| C 42 | 71.8 | 3.8 | 1566 | 12 | CG757757 | CG757757 P053-1-D0  |
| C 43 | 71.6 | 3.8 | 1578 | 10 | DV791374 | DV791374 Hw LoIn 1  |
| C 44 | 71.6 | 3.8 | 975  | 3  | BU960792 | BU960792 AGENCOURT  |
| C 45 | 71.6 | 3.8 | 1059 | 14 | CNS00228 | AL097133 Drosophila |

#### ALIGNMENTS

RESULT 1  
LOCUS B07882/c 500 bp DNA linear GSS 13-MAR-1998  
DEFINITION YX102T7 pluescript expression library of genomic DNA from OGI-R  
Enterococcus faecalis genomic clone YX102 similar to hypothetical protein 2 of Lactobacillus leichmannii (X81869), genomic survey sequence.

ACCESSION B07882  
VERSION B07882.1 GI:2058762  
KEYWORDS GSS.  
SOURCE Enterococcus faecalis  
ORGANISM Enterococcus faecalis  
Bacteria; Firmicutes; Lactobacillales; Enterococcaceae; Enterococcus.  
REFERENCE 1 (bases 1 to 500)  
AUTHORS Xu, Y., Jiang, L., Murray, B. E. and Weinstein, G. M.  
TITLE Enterococcus faecalis antigens in human infections  
JOURNAL Infect. Immun. 65 (10), 4207-4215 (1997)  
PUBMED 9317028  
COMMENT Contact: Weinstein, GM  
Department of Biochemistry and Molecular Biology  
University of Texas Medical School  
6431 Fannin Street, Houston, TX 77030, USA  
Tel: 713 500 6083  
Fax: 713 500 0652  
Email: georgew@utmsi.med.utmc.edu  
Seq primer: pluescript SK (-) c7  
Classes: shotgun  
High quality sequence stop: 500.  
Location/Qualifiers  
1..500  
/organism="Enterococcus faecalis"  
/mol\_type="genomic DNA"  
/strain="OGIR"  
/db\_xref="taxon:1351"  
/clone="YX102"  
/lab\_host="XLIBLue MRP"  
/clone\_lib="pluescript expression library of genomic DNA from OGI-R"  
/note="Vector: pluescript SK(-), Site 1: HindIII, Site 2: EcoRI; genomic DNA from an Enterococcus faecalis strain OGI-R was partially digested with SauAI, sized

#### FEATURES

source

fractionated and ligated into the BamH site of cosmid vectors pLARX and pBelobAC11. After in vitro packaging and plating, the cosmid clones were screened with antisera from 4 enterococcal endocarditis patients and a rabbit immunized with surface proteins from an E. faecalis isolate. Cosmid DNA from immunopositive cosmid clones was treated with Dnae I, the ends were filled with the Klenow fragment, and then ligated into the EcoRV site in pBluescript SK(-). The ligation mixture was transformed into XL1Blue MRF', and the transformants were screened with antisera. The immunopositive clones were selected for DNA sequencing."

## ORIGIN

Query Match 26.4%; Score 496.4; DB 11; Length 500;  
Best Local Similarity 99.4%; Pred. No. 6.2e-106;  
Matches 497; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

1319 AAAAGCAGATCTCTAGAAAGCTTTAAATAAGAGCGAATGTGATAAGGTCATACCG 1378  
|||||  
500 AAAAGCAGATCTCTAGAAAGCTTTAAATAAGAGCGAATGTGATAAGGTCATACCG 441  
|||||

1379 ACGACCAAAACACCCAACTGTTGAAGTTGACAGGTGGAAACGTTTCATTAAAGTCG 1438  
|||||  
440 ACGACCAAAACACCCAACTGTTGAAGTTGACAGGTGGAAACGTTTCATTAAAGTCG 381  
|||||

1439 ATGGGAGTGTACAGCGACACAGCCTTGCGGAGCTTCCTTGTGCGCGGTGATCAAA 1498  
|||||  
380 ATGGGAGTGTACAGCGACACAGCCTTGCGGAGCTTCCTTGTGCGCGGTGATCAAA 321  
|||||

1499 ACAGCGACACACCAAAATTATTTGAAATGATGAAACAAAGAGACAGCACTTGGGTGA 1558  
|||||  
320 ACAGCGACACACCAAAATTATTTGAAATGATGAAACAAAGAGACAGCACTTGGGTGA 261  
|||||

1559 AAACAAAGCTGAGACACTCTTTTACAAACGCGTGTGATGATTGATATACAG 1618  
|||||  
260 AAACAAAGCTGAGACACTCTTTTACAAACGCGTGTGATGATTGATATACAG 201  
|||||

1619 GGCTTAAATAGGGTCTCTTATTTAGAGAAAGCTAGCTCTGTGATGATTATGCTTGT 1678  
|||||  
200 GGCTTAAATAGGGTCTCTTATTTAGAGAAAGCTAGCTCTGTGATGATTATGCTTGT 141  
|||||

1679 TAACAAATCGATTGAAATTTGTGTCATGAAACAATCATATATGACAAAGAAACCTAG 1738  
|||||  
140 TAACAAATCGATTGAAATTTGTGTCATGAAACAATCATATATGACAAAGAAACCTAG 81  
|||||

1739 TTTCACGAGAAAAGTACCAAAACAAACAGTAAGTACTTACCTTCAACAGGTGGCAAG 1798  
|||||  
80 TTTCACGAGAAAAGTACCAAAACAAACAGTAAGTACTTACCTTCAACAGGTGGCAAG 21  
|||||

1799 GAATCTAGCTTACTTAGGA 1818  
|||||  
20 GAATCTAGCTTACTTAGGA 1  
|||||

RESULT 2 500 bp DNA 1linear GSS 13-MAR-1998  
B07890/c  
LOCUS  
DEFINITION YX98T7 pBluescript expression library of genomic DNA from OG1RF  
Enterococcus faecalis genomic clone YX98 similar to hypothetical  
protein 2 of Lactobacillus leichmannii (X81869), genomic survey  
sequence.  
ACCESSION B07890  
VERSION B07890.1 GI:2058780  
KEYWORDS GSS.  
SOURCE Enterococcus faecalis  
ORGANISM Enterococcus faecalis  
Bacteria; Firmicutes; Lactobacillales; Enterococcaceae;  
Enterococcus.  
REFERENCE 1 (bases 1 to 500)  
AUTHORS Xu, Y., Jiang, L., Murray, B.E. and Weinstock, G.M.  
TITLE Enterococcus faecalis antigens in human infections  
JOURNAL Infect. Immun. 65 (10), 4207-4215 (1997)

## PUBMED

9317028

Contact: Weinstock, GM

Department of Biochemistry and Molecular Biology

University of Texas Medical School

6431 Fannin Street, Houston, TX 77030, USA

Tel: 713 500 6083

Fax: 713 500 0652

Email: georgew@utmsf.med.utmc.edu

Seq primer: pBluescript SK (-) t7

Class: shotgun

High quality sequence stop: 500.

## FEATURES

## source

1..500

/organism="Enterococcus faecalis"

/mol\_type="genomic DNA"

/strain="OG1RF"

/db\_xref="taxon:1351"

/clone="YX98"

/lab\_host="XL1Blue MRF'"

/clone\_lib="pBluescript expression library of genomic DNA  
from OG1RF"

/note="Vector: pBluescript SK(-); Site\_1: HindIII; Site\_2:  
EcoRI; Genomic DNA from an Enterococcus faecalis strain  
OG1RF was partially digested with Sau3AI, sized  
fractionated and ligated into the BamH site of cosmid  
vectors pLARX and pBelobAC11. After in vitro packaging  
and plating, the cosmid clones were screened with antisera  
from 4 enterococcal endocarditis patients and a rabbit  
immunized with surface proteins from an E. faecalis  
isolate. Cosmid DNA from immunopositive cosmid clones was  
treated with Dnae I, the ends were filled with the Klenow  
fragment, and then ligated into the EcoRV site in  
pBluescript SK(-). The ligation mixture was transformed  
into XL1Blue MRF', and the transformants were screened  
with antisera. The immunopositive clones were selected for  
DNA sequencing."

## ORIGIN

Query Match 26.2%; Score 492.6; DB 11; Length 500;  
Best Local Similarity 99.0%; Pred. No. 4.9e-105;  
Matches 495; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

1190 TGATGCTCTGAAAATTATCAAGTACGAAACAAAGGCTTCACTGTGCGGTTA 1249  
|||||  
500 TGATGCTCTGAAAATTATCAAGTACGAAACAAAGGCTTCACTGTGCGGTTA 441  
|||||

1250 ATCCAGCGTATATTCCTAGCGTAAACACGAGGGGCACTAAATTCCTTACTTATAC 1309  
|||||  
440 ATCCAGCGTATATTCCTAGCGTAAACACGAGGGGCACTAAATTCCTTACTTATAC 381  
|||||

1310 ATTTAAATGAAAAGAGATCTCTACGAAAGCTTTAAATAAGGCGAATGTTGATACG 1369  
|||||  
380 ATTTAAATGAAAAGAGATCTCTACGAAAGCTTTAAATAAGGCGAATGTTGATACG 321  
|||||

1370 GTCATACGACGACCAACACCACTGTTGAAGTTGTGACAGGTGGAAAACGTTTCA 1429  
|||||  
320 GTCATACGACGACCAACACCACTGTTGAAGTTGTGACAGGTGGAAAACGTTTCA 261  
|||||

1430 TTTAAAGTGATGGCGATGTGACAGGACCAAGCCTTGGCGGAGCTTCTTGTGCTCC 1489  
|||||  
260 TTTAAAGTGATGGCGATGTGACAGGACCAAGCCTTGGCGGAGCTTCTTGTGCTCC 201  
|||||

1490 GTGATCAAAAACAGGACACAGCAATTTTGAATGCAATGAACAAAGAAAGACGAA 1549  
|||||  
200 GTGATCAAAAACAGGACACAGCAATTTTGAATGCAATGAACAAAGAAAGACGAA 141  
|||||

1550 CTTGGGTGAAAACAAAGCTGAGCACTATTTTACAAACAGGCTGATGATTAGTTG 1609  
|||||  
140 CTTGGGTGAAAACAAAGCTGAGCACTATTTTACAAACAGGCTGATGATTAGTTG 81  
|||||

1610 ATATACAGGGCTTAAATACGTTACTTATTTTGAAGAAAACGTAGTCTCTGATGATT 1669  
|||||  
80 ATATACAGGGCTTAAATACGTTACTTATTTTGAAGAAAACGTAGTCTCTGATGATT 21  
|||||



QY 1670 ATGCTGTGTAACAATCGG 1689  
DB 20 ATGCTGTGTAACAATCGG 1

RESULT 3  
B07866/c  
LOCUS  
DEFINITION B07866 500 bp DNA linear GSS 13-MAR-1998  
YX106T3 pBluescript expression library of genomic DNA from OG1RF  
Enterococcus faecalis genomic clone YX106 similar to hypothetical  
protein 2 of Lactobacillus leichmannii (X81869), genomic survey  
sequence.

ACCESSION B07866  
VERSION B07866.1 GI:2058766  
KEYWORDS GSS.  
SOURCE Enterococcus faecalis  
ORGANISM Enterococcus faecalis  
Bacteria; Firmicutes; Lactobacillales; Enterococcaceae;  
Enterococcus.

REFERENCE 1 (bases 1 to 500)  
AUTHORS Xu, Y., Jiang, L., Murray, B. E. and Weinstein, G. M.  
TITLE Enterococcus faecalis antigens in human infections  
JOURNAL Infect. Immun. 65 (10), 4207-4215 (1997)  
PUBMED 9317028

COMMENT Contact: Weinstein, G.M.  
Department of Biochemistry and Molecular Biology  
University of Texas Medical School  
6431 Fannin Street, Houston, TX 77030, USA  
Tel: 713 500 6083  
Fax: 713 500 0652  
Email: georgew@utmsg.med.utmc.edu  
Seg primer: pBluescript SK (-) c7  
Class: shotgun  
High quality sequence stop: 500.  
Location/Qualifiers  
1..500  
/organism="Enterococcus faecalis"  
/mol\_type="genomic DNA"  
/strain="OG1RF"  
/db\_xref="taxon:1351"  
/clone="YX106"  
/lab\_host="XL1Blue MRP"  
/clone\_1lb="pBluescript expression library of genomic DNA  
from OG1RF"  
/note="Vector: pBluescript SK(-); Site 1: HindIII; Site 2:  
EcoRI; Genomic DNA from an Enterococcus faecalis strain  
OG1RF was partially digested with Sau3AI, sized  
fractionated and ligated into the BamHI site of cosmid  
vectors pIAFRx and pReloBAC11. After in vitro packaging  
and plating, the cosmid clones were screened with antisera  
from 4 enterococcal endocarditis patients and a rabbit  
immunized with surface proteins from an E. faecalis  
isolate. Cosmid DNA from immunopositive cosmid clones was  
treated with Dnae I, the ends were filled with the Klenow  
fragment, and then ligated into the EcoRV site in  
pBluescript SK(-). The ligation mixture was transformed  
into XL1Blue MRP, and the transformants were screened  
with antisera. The immunopositive clones were selected for  
DNA sequencing."

ORIGIN  
Query Match 26.2%; Score 492; DB 11; Length 500;  
Best Local Similarity 99.0%; Pred. No. 6.8e-105;  
Matches 495; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 997 GGTAAAGATGTGGCAATTTGTAATATCAATTTCTGTAATATTCATG 1056  
DB 500 GGTAAAGATGTGGCAATTTGTAATATCAATTTCTGTAATATTCATG 441

QY 1057 GGGATTGCAGCAAGAGCGCATTAATAATAGTCAAAATTCATTAGTTATAA 1116  
DB 440 GGGATTGCAGCAAGAGCGCATTAATAATAGTCAAAATTCATTAGTTATAA 381

QY 1117 CATGATGAGCCTTAACCTTTGATTAACGTAAGCTTCGAGAGATGCTTATAT 1176  
DB 380 CATGATGAGCCTTAACCTTTGATTAACGTAAGCTTCGAGAGATGCTTATAT 321

QY 1177 GATGGGATACAGTATGCTCTCTGTAATTAATCAAGTACGTAAACAAATGCTTC 1236  
DB 320 GATGGGATACAGTATGCTCTCTGTAATTAATCAAGTACGTAAACAAATGCTTC 261

QY 1237 ACTGTGCGGTTAATTCACGATATATCTTACGCTTACACCGGCGGACACTAAATTC 1296  
DB 260 ACTGTGCGGTTAATTCACGATATATCTTACGCTTACACCGGCGGACACTAAATTC 201

QY 1297 GTTTACTTATGATTAATGAAAAGCAGATCTTACGAAAGGCTTTAAATGAGGG 1356  
DB 200 GTTTACTTATGATTAATGAAAAGCAGATCTTACGAAAGGCTTTAAATGAGGG 141

QY 1357 AATGTTGATTAACGTCATACCGACGACCAACCAACCACTGTTGAAGTTGACAGT 1416  
DB 140 AATGTTGATTAACGTCATACCGACGACCAACCAACCACTGTTGAAGTTGACAGT 81

QY 1417 GGGAAACGTTTCAATTAAGTCAGTGGCGCATGTCAGCAGCACACAGCTTGGCGGAGCT 1476  
DB 80 GGGAAACGTTTCAATTAAGTCAGTGGCGCATGTCAGCAGCACACAGCTTGGCGGAGCT 21

QY 1477 TCCTTGTGTCGCGTGATCA 1496  
DB 20 TCCTTGTGTCGCGTGATCA 1

RESULT 4  
B07865  
LOCUS  
DEFINITION YX106T3 pBluescript expression library of genomic DNA from OG1RF  
Enterococcus faecalis genomic clone YX106, genomic survey sequence.

ACCESSION B07865  
VERSION B07865.1 GI:2058765  
KEYWORDS GSS.  
SOURCE Enterococcus faecalis  
ORGANISM Enterococcus faecalis  
Bacteria; Firmicutes; Lactobacillales; Enterococcaceae;  
Enterococcus.

REFERENCE 1 (bases 1 to 450)  
AUTHORS Xu, Y., Jiang, L., Murray, B. E. and Weinstein, G. M.  
TITLE Enterococcus faecalis antigens in human infections  
JOURNAL Infect. Immun. 65 (10), 4207-4215 (1997)  
PUBMED 9317028

COMMENT Contact: Weinstein, G.M.  
Department of Biochemistry and Molecular Biology  
University of Texas Medical School  
6431 Fannin Street, Houston, TX 77030, USA  
Tel: 713 500 6083  
Fax: 713 500 0652  
Email: georgew@utmsg.med.utmc.edu  
Seg primer: pBluescript SK (-) c3  
Class: shotgun  
High quality sequence stop: 450.  
Location/Qualifiers  
1..450  
/organism="Enterococcus faecalis"  
/mol\_type="genomic DNA"  
/strain="OG1RF"  
/db\_xref="taxon:1351"  
/clone="YX106"  
/lab\_host="XL1Blue MRP"  
/clone\_1lb="pBluescript expression library of genomic DNA  
from OG1RF"  
/note="Vector: pBluescript SK(-); Site 1: HindIII; Site 2:  
EcoRI; Genomic DNA from an Enterococcus faecalis strain  
OG1RF was partially digested with Sau3AI, sized  
fractionated and ligated into the BamHI site of cosmid  
vectors pIAFRx and pReloBAC11. After in vitro packaging  
and plating, the cosmid clones were screened with antisera

from 4 enterococcal endocarditis patients and a rabbit immunized with surface proteins from an E. faecalis isolate. Cosmid DNA from immunopositive cosmid clones was treated with Dnae I, the ends were filled with the Klenow fragment, and then ligated into the EcoRV site in pBluescript SK(-). The ligation mixture was transformed into XL1Blue MRF, and the transformants were screened with antisera. The immunopositive clones were selected for DNA sequencing."

## ORIGIN

Query Match 23.1%; Score 434.8; DB 11; Length 450;  
Best Local Similarity 99.1%; Pred. No. 2e-91;  
Matches 447; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 725 AAGAACAGCAAAAGCGTTTATTACTGGGAAAAGTTATGAAATGGCGGAAATGATTTC 784  
DB 1 AAGAACAGCAAAAGCGTTTATTACTGGGAAAAGTTATGAAATGGCGGAAATGATTTC 60  
QY 785 CAGAGCAGAGAAATGGAACGGAGAAATTAAAGTTTAAATCTTGAGTGTGGTGTATA 844  
DB 61 CAGAGCAGAGAAATGGAACGGAGAAATTAAAGTTTAAATCTTGAGTGTGGTGTATA 120  
QY 845 TTTTGAAGAAAGTAAAGCTCCAAATTAATGAGAATTAAATGAAAATCAACAAAACAC 904  
DB 121 TTTTGAAGAAAGTAAAGCTCCAAATTAATGAGAATTAAATGAAAATCAACAAAACAC 180  
QY 905 CATTTACATTTAGCAAAACATCAACACCTGTGAAAAAAGCGTCAAAAATGATACCT 964  
DB 181 CATTTACATTTAGCAAAACATCAACACCTGTGAAAAAAGCGTCAAAAATGATACCT 240  
QY 965 CTAAGTGTATTAACAACACCAAGCTGTAGTGTAAAGATGGCAATGGCGAAAAA 1024  
DB 241 CTAAGTGTATTAACAACACCAAGCTGTAGTGTAAAGATGGCAATGGCGAAAAA 300  
QY 1025 TTTAAATATCAAAATTTCTGTAATATTTCCATTGGGGATTGCAAGAAAGCGAGCTA 1084  
DB 301 TTTAAATATCAAAATTTCTGTAATATTTCCATTGGGGATTGCAAGAAAGCGAGCTA 359  
QY 1085 ATTAATAGTCGAATTTCAATTTAGTGAATAAACATGATGACGCTTAATCTTTGATACG 1144  
DB 360 ATTAATAGTCGAATTTCAATTTAGTGAATAAACATGATGACGCTTAATCTTTGATACG 419  
QY 1145 TGACTTCTGAGAGATGCTTATGCGTTATA 1175  
DB 420 TGACTTCTGAGAGATGCTTATGCGTTATA 450

## RESULT 5

B07881 450 bp DNA linear GSS 13-MAR-1998  
LOCUS YX102T3 pBluescript expression library of genomic DNA from OG1RF  
DEFINITION Enterococcus faecalis genomic clone YX102, genomic survey sequence.  
ACCESSION B07881  
VERSION B07881.1 GI:2058761  
KEYWORDS GSS.  
SOURCE Enterococcus faecalis  
ORGANISM Enterococcus faecalis  
Bacteria; Firmicutes; Lactobacillales; Enterococcaceae;  
Enterococcus.  
1 (bases 1 to 450)  
Xu, Y., Jiang, L., Murray, B.E. and Weinstein, G.M.  
TITLE Enterococcus faecalis antigens in human infections  
JOURNAL Infect. Immun. 65 (10), 4207-4215 (1997)  
PUBMED 9317028  
COMMENT Contact: Weinstein, GM  
Department of Biochemistry and Molecular Biology  
University of Texas Medical School  
6431 Fannin Street, Houston, TX 77030, USA  
Tel: 713 500 6083  
Fax: 713 500 0652  
Email: georgew@utmsg.med.utah.tmc.edu  
Seq primer: pBluescript SK(-) c3

Class: shotgun  
High quality sequence stop: 450.

## FEATURES

source

1..450  
/organism="Enterococcus faecalis"  
/mol\_type="genomic DNA"  
/strain="OG1RF"  
/db\_xref="taxon:1351"  
/clone="YX102"  
/lab\_host="XL1Blue MRF"  
/clone\_11b="pBluescript expression library of genomic DNA from OG1RF"  
/note="Vector: pBluescript SK(-); Site 1: HindIII; Site 2: EcoRI; Genomic DNA from an Enterococcus faecalis strain OG1RF was partially digested with Sau3AI, sized fractionated and ligated into the BamHI site of cosmid vectors pLAFR and pReloBAC11. After in vitro packaging and plating, the cosmid clones were screened with antisera from 4 enterococcal endocarditis patients and a rabbit immunized with surface proteins from an E. faecalis isolate. Cosmid DNA from immunopositive cosmid clones was treated with Dnae I, the ends were filled with the Klenow fragment, and then ligated into the EcoRV site in pBluescript SK(-). The ligation mixture was transformed into XL1Blue MRF, and the transformants were screened with antisera. The immunopositive clones were selected for DNA sequencing."

## ORIGIN

Query Match 22.9%; Score 430.2; DB 11; Length 450;  
Best Local Similarity 98.2%; Pred. No. 2.5e-90;  
Matches 435; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1070 AAGAGGAGCGCTTAATTAATCGTCAATTTAGTTATGAATGAATGACGCT 1129  
DB 7 AAGAGGAGCGCTTAATTAATCGTCAATTTAGTTATGAATGAATGACGCT 66  
QY 1130 TAACTTTGATPAACGCTTCTGAGAGATGCTTAAGCGTTATATGATGAGGATACAG 1189  
DB 67 TAACTTTGATPAACGCTTCTGAGAGATGCTTAAGCGTTATATGATGAGGATACAG 126  
QY 1190 TGATTCCTCTGAAATTAATCAAGTACTGAACAAGCAATGGCTTCACTGCGCGTTA 1249  
DB 127 TGATTCCTCTGAAATTAATCAAGTACTGAACAAGCAATGGCTTCACTGCGCGTTA 186  
QY 1250 ATCCAGCGTATTTCTTAAGCTTAACACGAGGCGGACACTAAATTCGTTACTTATGC 1309  
DB 187 ATCCAGCGTATTTCTTAAGCTTAACACGAGGCGGACACTAAATTCGTTACTTATGC 246  
QY 1310 ATTTAATGAATAAAGCAGATCCTACGAAGGCTTTAAATAGGCGAATGTGATACG 1369  
DB 247 ATTTAATGAATAAAGCAGATCCTACGAAGGCTTTAAATAGGCGAATGTGATACG 306  
QY 1370 GTCATACGACGACCAACACCACTGTTAAAGTTGTGACAGGTGGAACGTTTCA 1429  
DB 307 GTCATACGACGACCAACACCACTGTTAAAGTTGTGACAGGTGGAACGTTTCA 366  
QY 1430 TTTAAGTGTATGCGATGTGACAGGACCAAGCCTTGGCGGAGCTTCTTTGTCGTC 1489  
DB 367 TTTAAGTGTATGCGATGTGACAGGACCAAGCCTTGGCGGAGCTTCTTTGTCGTC 426  
QY 1490 GTGATCAAAACGACGACAGCA 1512  
DB 427 GTGATCAAAACGACGACAGCA 449

## COMMENT

RESULT 6  
B07889 400 bp DNA linear GSS 13-MAR-1998  
LOCUS YX98T3 pBluescript expression library of genomic DNA from OG1RF  
DEFINITION Enterococcus faecalis genomic clone YX98, genomic survey sequence.  
ACCESSION B07889  
VERSION B07889.1 GI:2058779

KEYWORDS GSS.  
SOURCE Enterococcus faecalis  
ORGANISM Enterococcus faecalis  
Bacteria; Firmicutes; Lactobacillales; Enterococcaceae;  
Enterococcus.  
REFERENCE 1 (bases 1 to 400)  
AUTHORS Xu, Y., Jiang, L., Murray, B.E. and Weinstein, G.M.  
TITLE Enterococcus faecalis antigens in human infections  
JOURNAL Infect. Immun. 65 (10), 4207-4215 (1997)  
PUBMED 9317028  
COMMENT Contact: Weinstein, GM  
Department of Biochemistry and Molecular Biology  
University of Texas Medical School  
6431 Fannin Street, Houston, TX 77030, USA  
Tel: 713 500 6083  
Fax: 713 500 0652  
Email: georgew@utmsi.med.utah.edu  
Seq primer: pBluescript SK (-) c3  
Class: shotgun  
High quality sequence stop: 400.  
Location/Qualifiers  
1..400  
/organism="Enterococcus faecalis"  
/mol\_type="genomic DNA"  
/strain="OG1RF"  
/db\_xref="taxon:1351"  
/clone="YX98"  
/lab\_host="XL1Blue MRP"  
/clone\_1fb="pBluescript expression library of genomic DNA  
from OG1RF"  
/note="Vector: pBluescript SK(-); Site 1: HindIII; Site 2:  
EcoRI; Genomic DNA from an Enterococcus faecalis strain  
OG1RF was partially digested with Sau3AI, sized  
fractionated and ligated into the BamHI site of cosmid  
vectors pLARX and pBel0AC11. After in vitro packaging  
and plating, the cosmid clones were screened with antisera  
from 4 enterococcal endocarditis patients and a rabbit  
immunized with surface proteins from an E. faecalis  
isolate. Cosmid DNA from immunopositive cosmid clones was  
treated with Dnaase I, the ends were filled with the Klenow  
fragment, and then ligated into the EcoRV site in  
pBluescript SK(-). The ligation mixture was transformed  
into XL1Blue MRP, and the transformants were screened  
with antisera. The immunopositive clones were selected for  
DNA sequencing."

## ORIGIN

Query Match 20.2%; Score 380; DB 11; Length 400;  
Best Local Similarity 98.3%; Pred. No. 1.7e-78;  
Matches 394; Conservative 0; Mismatches 6; Indels 1; Gaps 1;  
QY 1032 TCMAATTTGTGAATATTCCTCATGGGATTGACAGCAAGAGCGGCTTAAATA 1091  
DB 1 TCMAATTTGTGAATATTCCTCATGGGATTGACAGCAAGAGCGGCTTAAATA 59  
QY 1092 GGTCAATTCATTTAGTGTGAATATTCCTCATGGGATTGACAGCAAGAGCGGCTT 1151  
DB 60 GGTCAATTCATTTAGTGTGAATATTCCTCATGGGATTGACAGCAAGAGCGGCTT 119  
QY 1152 TGGAGATGATGCTTATGCTTATGATGGGATACAGATGATGCTCTGAAATATTA 1211  
DB 120 TGGAGATGATGCTTATGCTTATGATGGGATACAGATGATGCTCTGAAATATTA 179  
QY 1212 AGTGAAGCAAGCAATGCGCTTCACTGCGCGCTTAATCCAGGCTAATTTCTTACGCT 1271  
DB 180 AGTGAAGCAAGCAATGCGCTTCACTGCGCGCTTAATCCAGGCTAATTTCTTACGCT 239  
QY 1272 AACACAGGCGGACACTAAATTTGTTACTTATGATTTAAATGAAGAGCAGATCC 1331  
DB 240 AACGCGAGGCGGACACTAAATTTGTTACTTATGATTTAAATGAAGAGCAGATCC 299  
QY 1332 TACGAAGGCTTTAAATGAAGCGATGTTATACGTCATCCGACGACCAACACC 1391  
|||||

DB 300 TACGAAGGCTTTAAATGAAGCGATGTTATACGTCATCCGACGACCAACACC 359  
QY 1392 ACCAAGCTTTGAAGTGTGACAGGTGGGAAGCTTTCATTA 1432  
DB 360 ACCAAGCTTTGAAGTGTGACAGGTGGGAAGCTTTCATTA 400  
|||||

RESULT 7  
LOCUS A2548467  
DEFINITION Entamoeba histolytica Sheared DNA Entamoeba histolytica  
genomic, genomic survey sequence.  
ACCESSION A2548467  
VERSION A2548467.1 GI:11172102  
KEYWORDS GSS.  
SOURCE Entamoeba histolytica  
ORGANISM Entamoeba histolytica  
Eukaryota; Entamoebidae; Entamoeba.  
REFERENCE 1 (bases 1 to 908)  
AUTHORS Loftus, B., Van Aken, S. and Fraser, C.  
TITLE Determination of clone end sequences from Entamoeba histolytica  
HMI:IMSS sheared DNA library  
JOURNAL Unpublished (2000)  
COMMENT Contact: Brendan J Loftus  
Department of Eukaryotic Genomics  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850, USA  
Tel: 301 838 0208  
Fax: 301 838 3543  
Email: b.loftus@tigr.org  
Clones are derived from the Entamoeba histolytica HMI:IMSS sheared  
DNA library  
Seq primer: M13-Reverse  
Classes: Shotgun  
High quality sequence start: 17  
High quality sequence stop: 828.  
Location/Qualifiers  
1..908  
/organism="Entamoeba histolytica"  
/mol\_type="genomic DNA"  
/strain="HMI:IMSS"  
/db\_xref="taxon:5759"  
/clone\_1fb="Entamoeba histolytica Sheared DNA"  
/note="Vector: PHOS1; Site 1: Bat I; Constructed at The  
Institute for Genomic Research (TIGR), Rockville, MD.  
Genomic DNA isolated from broth cultures of E. histolytica  
using a method described by Clark and Diamond (Clark,  
C.G., and Diamond, L.S. (1993) Entamoeba histolytica: a  
method for isolate identification. Exp. Parasitol.  
77:450.). The DNA was mechanically sheared to give a  
tight size distribution (~2 kb). The v + 1 method used for  
the library construction is described in detail in Smith,  
H.O. and Venter, J.C. (Making small insert libraries for  
whole genome shotgun sequencing projects. in Genome  
Sequencing: A Practical Approach, eds. M. Vaubin and B.  
Barrell, Oxford University Press, 1999)."

## FEATURES

## source

Query Match 4.8%; Score 91.2; DB 11; Length 908;  
Best Local Similarity 44.4%; Pred. No. 2.4e-10;  
Matches 366; Conservative 0; Mismatches 458; Indels 0; Gaps 0;  
QY 324 TGGGAACCTGTTGCTCAAGGAACCAACCGATGCAATGGAGTGTACGTTCAGTTAC 383  
DB 45 TATATGAAGAGTATGATGAAGAAAGATATATGAAGAGATGATGATGAAGAGA 104  
QY 384 TAAAAAACAATATGTTAAAGATGACGTATATCAATTAAGAGAACCAAAAGAGGTGT 443  
DB 105 TATATGAAGAGTATGATGAAGAAAGATATATGAAGATGATGAAGAGAGATGA 164  
QY 444 AGTGTGCTACGAATATGTGTGCGCTTCCAGATTTTCGAATGAATGATCAAGCAACAA 503  
DB 165 AGACGACGAATATGATTAAGAGATGATGATGAAGAAAGAGATGATGATGAAGA 224  
|||||

## ORIGIN

| QY         | 504                                                                |                       | CGGTCCTCTAATAATATGCAACGAAGAATTACGGTGTTCATATTTATCTTAAAAATGT   | 563    |
|------------|--------------------------------------------------------------------|-----------------------|--------------------------------------------------------------|--------|
| Db         | 225                                                                |                       | TGATGAAGACGATGGAAGACGACGAATATGAATTTAAGATGATGATGATGAAGAAAGA   | 284    |
| QY         | 564                                                                |                       | GGTACCCATGATGCTAGCTTTTACATGTGAAAAAATAGAACTGCTGAAAAATGAAGATT  | 623    |
| Db         | 285                                                                |                       | AGATGATGATGAAGAAGAAAGATGATGATGAAGATGATGAAGATGATGAAGAAAGATGA  | 344    |
| QY         | 624                                                                |                       | AAATGGCGCAATTTGTTTATTTCTPAAAAGCGAAGCTTCACGAGCAACGTAAATATAT   | 683    |
| Db         | 345                                                                |                       | AGATGATGAAGAAGATGATGAAGATGATGAAGATGAAGAAGAAAGATGATGAAGATGA   | 404    |
| QY         | 684                                                                |                       | CCAAAGATCAAAGATGATTTATATCATGCAACACGATTAAGAACAAGCAAAACGCTT    | 743    |
| Db         | 405                                                                |                       | TGAACACATGGAAGACGACGAATATGAATTTAGAAATGATGATGATGAAGAAAGATGA   | 464    |
| QY         | 744                                                                |                       | TATTACTGGGAAAAGTTATGAAATTTGGCGAAAATGATTTACAGAAAGCAGAGATGGAAC | 803    |
| Db         | 465                                                                |                       | TGATGAAGAAGAAAGATGATGATGAAGATGATGATGAAGAAAGATGATGAAGATGA     | 524    |
| QY         | 804                                                                |                       | GGAGATTTAACGTTAAAAATCTTGAAGTTGTTCTATATTTTGAAGAAGTAAAAAC      | 863    |
| Db         | 525                                                                |                       | TGAAGAAAGATGGAAGATGATGATGAAGAAAGAAATGATGATGAAGAAAGATGAAGA    | 584    |
| QY         | 864                                                                |                       | TCCAATTAATGCAGAATTAATGGAATAATCAACAAACACCATTTACATTTGAAGCAA    | 923    |
| Db         | 585                                                                |                       | AGAAAGATGATGGAAGAAAGATGGAAGAAAGAAAGAAAGATGATGATGAAGACGACGAAT | 644    |
| QY         | 924                                                                |                       | CAATCAAAACCGTTGTGAAAAAAACAGTCAAAATATATACCTTAAGTTGATPAAAAAC   | 983    |
| Db         | 645                                                                |                       | TGAATTAAGAAAGATGATGAAGAAGATGATGAAGAAGATGATGAAGAAGATGAAGA     | 704    |
| QY         | 984                                                                |                       | ACCAAGCTTAAGTGTGAAGATGTGGCAATTGGCGAAAAATTAATATCAAAATTTCTGT   | 1043   |
| Db         | 705                                                                |                       | AGAAAGATGATGATGATGATGATGAAGAAAGATGATGATGAAGATGATGAAGAAAGATGA | 764    |
| QY         | 1044                                                               |                       | AAATTTCCATTTGGGATTTGCGACAAAGAAAGCGACCTTAATATCGTCAATTTCA      | 1103   |
| Db         | 765                                                                |                       | AGATGATGAAGATGATGAAGAAGATGAAGATGATGAAGATGATGAAGATGATGAAGATGA | 824    |
| QY         | 1104                                                               |                       | TTTAGTTGATTAACATGATGCGACCTTAACCTTTGATTAACGTGA                | 1147   |
| Db         | 825                                                                |                       | TGAAGATGATGAAGACGACGAATTTGAATTTAAAGATGATGA                   | 868    |
| RESULT 8   |                                                                    |                       |                                                              |        |
| AZ551618/c |                                                                    |                       |                                                              |        |
| LOCUS      | AZ551618                                                           | 843 bp                | DNA                                                          | linear |
| DEFINITION | ENTR154STR Entamoeba histolytica Sheared DNA                       | Entamoeba histolytica |                                                              |        |
| ACCESSION  | AZ551618                                                           |                       |                                                              |        |
| VERSION    | AZ551618.1                                                         | GI:11176919           |                                                              |        |
| KEYWORDS   | GSS.                                                               |                       |                                                              |        |
| SOURCE     | Entamoeba histolytica                                              |                       |                                                              |        |
| ORGANISM   | Entamoeba histolytica                                              |                       |                                                              |        |
| REFERENCE  | 1 (bases 1 to 843)                                                 |                       |                                                              |        |
| AUTHORS    | Loftus,B., Van Aken,S. and Fraser,C.                               |                       |                                                              |        |
| TITLE      | Determination of clone end sequences from Entamoeba histolytica    |                       |                                                              |        |
| JOURNAL    | HMI:IMSS sheared DNA library                                       |                       |                                                              |        |
| COMMENT    | Unpublished (2000)                                                 |                       |                                                              |        |
|            | Contact: Brendan J Loftus                                          |                       |                                                              |        |
|            | Department of Eukaryotic Genomics                                  |                       |                                                              |        |
|            | The Institute for Genomic Research                                 |                       |                                                              |        |
|            | 9712 Medical Center Dr., Rockville, MD 20850, USA                  |                       |                                                              |        |
|            | Tel: 301 838 0208                                                  |                       |                                                              |        |
|            | Fax: 301 838 3543                                                  |                       |                                                              |        |
|            | Email: b1loftus@ligr.org                                           |                       |                                                              |        |
|            | Clones are derived from the Entamoeba histolytica HMI:IMSS sheared |                       |                                                              |        |
|            | DNA library                                                        |                       |                                                              |        |
|            | Seq primer: M13-Reverse                                            |                       |                                                              |        |
|            | Class: shotgun                                                     |                       |                                                              |        |

```

FEATURES
    source
        High quality sequence start: 39
        High quality sequence stop: 838.
        Location/Qualifiers
            1. 843
                /organism="Entamoeba histolytica"
                /mol_type="genomic DNA"
                /strain="HMI-1MS8"
                /db_xref="taxon:5759"
                /clone_id="Entamoeba histolytica Sheared DNA"
                /note="Vector: pHOsi; Site 1: Bat I; Constructed at The
                Institute for Genomic Research (TIGR), Rockville, MD.
                Genomic DNA isolated from broth cultures of B. histolytica
                using a method described by Clark and Diamond (Clark,
                C.G., and Diamond, L.S. (1993) Entamoeba histolytica: a
                method for isolate identification. Exp. Parasitol.
                77:450.) The DNA was mechanically sheared to give a
                tight size distribution (~2 kb). The v + i method used for
                the library construction is described in detail in Smith,
                H.O. and Venter, J.C. (Making small insert libraries for
                whole genome shotgun sequencing projects. In Genome
                Sequencing: A Practical Approach, eds. M. Vaudin and B.
                Barrell, Oxford University Press, 1993)."

```

## ORIGIN

| Query Match           | 4.7%;           | Score 88.8;        | DB 11;          | Length 843;                             |
|-----------------------|-----------------|--------------------|-----------------|-----------------------------------------|
| Best Local Similarity | 44.5%;          | Pred. No. 8.7e-10; |                 |                                         |
| Matches 351;          | Conservative 0; | Mismatches 437;    | Indels 0;       | Gaps 0;                                 |
| Qy                    | 335             | TTGCTCAAGAACCA     | CCGATGTC        | CAATTTGGGAATGTCATCTGTTCACTTAAACCAAA 394 |
| Db                    | 822             | TTGATGATGAAGAA     | GATGATGAAGAA    | GATGATGATGATGAAGAAAGATGATGAGAG 763      |
| Qy                    | 395             | ATGGTAAAGATGC      | AGTGTATACATTAA  | AGAAAGCCAAAGAGGGTGTAGTTGCTGTA 454       |
| Db                    | 762             | ATGATGAAGAAAG      | AGATGATGAAGATGA | AGATGATGAAGACGATGAAGACGACGAAT 703       |
| Qy                    | 455             | CGAATATGCTGTG      | TGGCTTCCCACTTAC | CAAAATGATCAAGCAAAACAGATGTTCTATA 514     |
| Db                    | 702             | ATGAAATTAAGAA      | GTATGATGATGAAG  | AAAGAAAGATGATGAAGATGAAGAACG 643         |
| Qy                    | 515             | AATATGGAACAGAA     | ATTAGCGGTGTTGAT | TTTATCCTTAAATTTGTGTAGCCAAATG 574        |
| Db                    | 642             | ATGAAACGACGA       | ATATGAAATTAAGAA | GATGATGATGATGAAGAAAGATGATGATG 583       |
| Qy                    | 575             | ATGCTACTTTACAT     | GTGCAAAAAAGTAG  | CACTCTGAAATGAAGATTTAAATGGCGCAG 634      |
| Db                    | 582             | AAGAAAGAAAGAT      | GATGATGAAGATGAT | GTAAAGAAAGATGAAGAAAGATGATGAAG 523       |
| Qy                    | 635             | AAATTTGTTATTT      | CTAAAGCGAAAGCT  | CACAGGCAACGTAAATATATCAACGAGTCA 694      |
| Db                    | 522             | AAAGTATGTAAG       | ATGATGATGAAGAA  | AGAAAGATGATGTAAGATGATGAAGACGATG 463     |
| Qy                    | 695             | AAAGTGAATTTAT      | PACTGSAACA      | CGATTAAGAACAGAAACGCTTTATTAATCGGA 754    |
| Db                    | 462             | AAAGCGACGAAT       | TGAATTTGAAGATG  | ATGATGAAGAAAGATGATGAAGAAAG 403          |
| Qy                    | 755             | AAAGTTTGAATTT      | GGCGAAATGATTT   | TCACAGAGCAGAGATYGGAAACGGGAAATTA 814     |
| Db                    | 402             | AAAGATGATGAT       | GAAGATGATGAAG   | AAAGATGATGAAGATGATGAAGAAAG 343          |
| Qy                    | 815             | CAGTTAAAAATCT      | TGAAGTTGGTTCG   | TATATTTAGAGAGATTAAGCTCCAAATATG 874      |
| Db                    | 342             | ATGAAGATGATG       | ATGAAGAAAGATG   | ATGATGAAGAAAGATGATGAAGAAAGATG 283       |
| Qy                    | 875             | CAGATTAATTTGA      | AAATCAAAACAAAC  | CACTTACAAATTGACCAACATCAAAAC 934         |
| Db                    | 282             | ATGAAGAAAGTAT      | GAAGAAAGATGAT   | ATGAAGACGACGAATATGATTTGAG 223           |
| Qy                    | 935             | CTGTGAAAAAACA      | GTCAAAATATGAT   | ACTTAAAGTGTATTAACAAACCAACGCTTAG 994     |
| Db                    | 222             | ATGATGAAGAA        | GATGATGAAGAA    | GATGATGAAGAAAGATGATGAAGAAAGATG 163      |
| Qy                    | 995             | ATGGTAAAGATG       | TGGCAATTTGGCA   | AAAAATTAATCAATTTCTGTAAATATTCAT 1054     |

```

Db      162 ATGATGATGATGATGAAGATGATGATGAAGATGATGAAGATGATGAAG 103
Qy      1055 TGGGATTCGACACAAAGAGCGACGCTAATTAATACCTCAATTCATTTAGTTGATA 1114
Db      102 AAGATGAAGATGATGAAGATGATGAAGATGATGAAGATGATGAAGATGATGAAG 43
Qy      1115 AACATGAT 1122
Db      42 AATATTAAT 35

RESULT 9
CNS06PLB/c 1084 bp DNA linear GSS 05-JUL-2001
LOCUS T7 end of clone AV0AA015D02 of library AV0AA from strain CBS 379 of
DEFINITION Saccharomyces exiguus, genomic survey sequence.
ACCESSION AL409509
VERSION AL409509.1 GI:12177058
KEYWORDS GSS.
SOURCE Saccharomyces exiguus
ORGANISM Saccharomyces exiguus
REFERENCE Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
AUTHORS Saccharomycetales; Saccharomycetaceae; Saccharomyces.
1 (bases 1 to 1084)
Bouclier,J.L., Aigle,M., Artiguenave,F., Blandin,G.,
de-Montigny,J., Dujon,B., Durren,P., Lepingle,A., Llorente,B.,
Malpertuy,A., Neuvéglise,C., Ozier-Kalogeropoulos,O., Potier,S.,
Saurin,M., Tekala,F., Toffano-Nicoche,C., Wesolowski-Louvel,M.,
Wincker,P. and Weissenbach,J.
Genomic exploration of the hemiascomycetous yeasts: 1. A set of
yeast species for molecular evolution studies
JOURNAL FEBS Lett. 487 (1), 3-12 (2000)
PUBMED 11152876
REFERENCE 2 (bases 1 to 1084)
AUTHORS Bon,E., Neuvéglise,C., Lepingle,A., Wincker,P., Artiguenave,F.,
Galliardin,C. and Casaregola,S.
Genomic exploration of the hemiascomycetous yeasts: 6.
Saccharomyces exiguus
JOURNAL FEBS Lett. 487 (1), 42-46 (2000)
PUBMED 11152881
REFERENCE 3 (bases 1 to 1084)
AUTHORS Genoscope.
Direct Submission
Submitted (07-SEP-2000) Genoscope - Centre National de Séquençage,
2 rue Gaston Crémieux, CP 5706, 91057 EVRY cedex, FRANCE. (E-mail :
seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
This GSS is part of a random genomic sequencing program of thirteen
yeast species: Saccharomyces bayanus var. uvarum, Saccharomyces
exiguus, Saccharomyces servazzii, Zygosaccharomyces rouxii,
Saccharomyces kluyveri, Kluyveromyces thermotolerans, Kluyveromyces
lactis var. lactis, Kluyveromyces marxianus var. marxianus, Pichia
angusta, Debaryomyces hansenii var. hansenii, Pichia sorbolicophila,
Candida tropicalis and Yarrowia lipolytica. Genomic inserts of 3 to
5 kb were prepared and both extremities were sequenced. See
keywords for description of this sequence and for the sequence of
the other extremity of this insert.
Location/Qualifiers
1..1084
/organism="Saccharomyces exiguus"
/mol_type="genomic DNA"
/strain="CBS 379"
/db_xref="taxon:34358"
/clone="AV0AA015D02"
/clone_1ib="AV0AA"
/notes="end : 17"
misc_feature
/inference="non-experimental evidence, no additional
details recorded"
/notes="similar to Saccharomyces cerevisiae ORF YMR304w" [
UBP15 ; similarity to human ubiquitin-specific protease ]
1 putative frameshift(s) "

```

```

ORIGIN
Query Match 4.2%; Score 79; DB 14; Length 1084;
Best Local Similarity 45.2%; Pred. No. 1.9e-07;
Matches 214; Conservative 17; Mismatches 242; Indels 0; Gaps 0;

Qy      670 ACGTAAATATATATCCAGAGAGTCAAGATGATTAATATACATGCAACGATTAAGAA 729
Db      556 AAAAAAAAAATTAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAATTAATAAAAAA 497
Qy      730 CAGCAAAACGCTTATTAATCTGGGAAAGATTATGAAATTCGGGAAATATCTTACAGAA 789
Db      436 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAATTAATAAAAA 437
Qy      790 GCAGGAATGCAAGCGGGAATTAACAGTTAAATTAATCTTGAGTTGCTCGATATTTTA 849
Db      436 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 377
Qy      850 GAAAGATTAAGTCCAAATTAATGTCAGATTAATTAATGAAATCAACAAACACCATTT 909
Db      376 GAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 317
Qy      910 ACAATTGAAGCAACATCAAAACACCTGTTGAAAAACAGTCAAAATGATACCTTAA 969
Db      316 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 257
Qy      970 GTTGATTAACACACCAAGCTTAGATGTTAAGATGTCGCAATTGCGGAAAAATTA 1029
Db      256 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 197
Qy      1030 TATCAATTTCTGTAAATATTCATTTGGGATTTGACGACAAAGAGCGACCTAATA 1089
Db      196 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 137
Qy      1090 TACGTCAATTCATTTAGTTGATTAACATGATGACGCTTACTTGTATTA 1142
Db      136 GARGAAGATTAATGATTAATAATGTTCTAATTCGAMATTAATTAATTTGATTA 84

RESULT 10
CG753732 1811 bp DNA linear GSS 24-OCT-2003
LOCUS CG753732/c
DEFINITION P048-4-G03.ya Ppa EcORI BAC library Pristionchus pacificus genomic,
genomic survey sequence.
ACCESSION CG753732
VERSION CG753732.1 GI:37978509
KEYWORDS GSS.
SOURCE Pristionchus pacificus
ORGANISM Pristionchus pacificus
REFERENCE Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida;
AUTHORS Neodiplogasteridae; Pristionchus.
1 (bases 1 to 1811)
Srinivasan,J., Sins,W., Jeeze,T., Wiggers-Perebolte,L., Jansen,K.,
Bunjer,J., van der Meulen,M. and Sommer,R.J.
An integrated physical and genetic map of the nematode Pristionchus
pacificus
JOURNAL Mol. Genet. Genomics 269 (5), 715-722 (2003)
PUBMED 12884007
COMMENT Contact: Sommer RJ
Evolutionary Biology
Max-Planck-Institute for Developmental Biology
Spemannstr. 37-39, Tuebingen D-72076, Germany
Tel: 00497071601371
Fax: 00497071601498
Email: ralf.sommer@uebingen.mpg.de
Class: BAC ends.
Location/Qualifiers
1..1811
/organism="Pristionchus pacificus"
/mol_type="genomic DNA"
/strain="California"
/db_xref="taxon:54126"
/clone_1ib="Ppa EcORI BAC library"

```

ORIGIN

/note="The library was generated by a partial digest of the genomic DNA with EcoRI and cloning into the BAC vector."

Query Match 4.2%; Score 78.2; DB 12; Length 1811;  
Best Local Similarity 37.8%; Pred. No. 3.2e-07;  
Matches 248; Conservative 0; Mismatches 408; Indels 0; Gaps 0;  
QY 381 ACTTAAACAAATGCTAAAGATGCACTGATATCCATTAAAGAAACCAAGAGGG 440  
DB 1712 ANNAANANANANANANANANANANANANANANANANANANANANANANANANA 1653  
QY 441 TGTAGTCTGCTAGCAATATGCTGTCGCTCCAGTTTCGAAATGATCAACCAAC 500  
DB 1652 ANNAANANANANANANANANANANANANANANANANANANANANANANANANA 1593  
QY 501 AGATGCTTCTATATATATGCAACAGAAATTCAGCGTGTTCATATTTATCTTAAAA 560  
DB 1592 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1533  
QY 561 TGTGTACCCATGATGCTGTTTACATGTGAAAAAGTAGAAGTCTGAAATGAAAG 620  
DB 1532 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1473  
QY 621 ATTAAATGCGCAGAAATTTGTTATTTCTTAAAGCGAAGCTCACAGGACAGTAAATA 680  
DB 1472 ANNAANANANANANANANANANANANANANANANANANANANANANANANANA 1413  
QY 681 TATCCAGAGTCAAGATGATTTATATCATGACCAACGATTAAGAAACAGCAAAACG 740  
DB 1412 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1353  
QY 741 CTTATTTCTGGAAGAGTTATGAAATTTGCGAAGATGTTTCAAGAACAGAGATGG 800  
DB 1352 NNANANANANANANANANANANANANANANANANANANANANANANANANA 1293  
QY 801 AACCGAGATTTAACAATTTAAATCTTGAGTGTGTTATTTTGAAGAAAGTAA 860  
DB 1292 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1233  
QY 861 AGCTCAATATATGCAATTTATGAAATCAACAAACACCACTTTACATTTGAAGC 920  
DB 1232 NAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1173  
QY 921 AAACATCAACACCTGTTGAAAAACAGTCAAAATGATCTCTTAAAGTTGATTAAC 980  
DB 1172 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1113  
QY 981 AACCCAGCTTGAATGTTAAAGATGTGCAATTTGGCAAAATTTAAATATCAAA 1036  
DB 1112 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1057

RESULT 11  
AZ531291/c 877 bp DNA linear GSS 03-NOV-2000

LOCUS AZ531291 Entamoeba histolytica Sheared DNA Entamoeba histolytica  
DEFINITION ENTBQ34TR genomic survey sequence.

ACCESSION AZ531291  
VERSION AZ531291.1 GI:11085838

KEYWORDS GSS.

SOURCE Entamoeba histolytica

ORGANISM Entamoeba histolytica

REFERENCE 1 (bases 1 to 877)  
AUTHORS Loftus, B., Van Aken, S. and Fraser, C.  
TITLE Determination of clone end sequences from Entamoeba histolytica

JOURNAL HMJ:IMSS sheared DNA library

COMMENT Unpublished (2000)  
Contact: Brendan J Loftus  
Department of Eukaryotic Genomics  
The Institute for Genomic Research

FEATURES

source

9712 Medical Center Dr., Rockville, MD 20850, USA  
Tel: 301 838 0208  
Fax: 301 838 3543  
Email: bjoftus@igf.org  
Clones are derived from the Entamoeba histolytica HMJ:IMSS sheared DNA library  
Seq primer: M13-Reverse  
Class: shotgun  
High quality sequence start: 22  
High quality sequence stop: 829.  
location/Qualifiers  
1..877  
/organism="Entamoeba histolytica"  
/mol\_type="genomic DNA"  
/strain="HMJ:IMSS"  
/db\_xref="taxon:5759"  
/clone\_id="Entamoeba histolytica Sheared DNA"  
/note="Vector: pHS1; Site 1: Bst I; Constructed at the Institute for Genomic Research (TIGR), Rockville, MD. Genomic DNA isolated from broth cultures of E. histolytica using a method described by Clark and Diamond (Clark, C.G., and Diamond, L.S. (1993) Entamoeba histolytica: a method for isolate identification. Exp. Parasitol. 77:450.). The DNA was mechanically sheared to give a tight size distribution (~2 kb). The v + i method used for the library construction is described in detail in Smith, H.O. and Venter, J.C. (Making small insert libraries for whole genome shotgun sequencing projects. in Genome Sequencing: A Practical Approach, eds. M. Vaubin and B. Barrell, Oxford University Press, 1999)."

ORIGIN

Query Match 4.1%; Score 78; DB 11; Length 877;  
Best Local Similarity 45.2%; Pred. No. 3.1e-07;  
Matches 328; Conservative 0; Mismatches 395; Indels 3; Gaps 1;  
QY 251 ACGAATTTACGAGCAAGCGGCGAGCGGATGATGACCTTAAACGCTGCC 310  
DB 805 ACGATATGATTTAGAAAGAAAGAGATGATGATGATGATGATGATGATGATGATG 746  
QY 311 AAAGTTAACTCTGGAACCTGTTCTCAAGAACCAACGATCAAAATGGAAATGCA 370  
DB 745 AAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 686  
QY 371 CTGTCACTTACCTTAAACAAATGCTAAAGATGCTGATATCCATTAAAGAAAGC 430  
DB 685 ATGAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 626  
QY 431 CAAAAGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 490  
DB 625 AAGACGATGAAGACGACGATATGATGATGATGATGATGATGATGATGATGATG 566  
QY 491 TCAACCAACAGATGTTCTATTAATATGGAACAGAAATTTAGCGTTGTCATATT 550  
DB 565 ATGATGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 506  
QY 551 ATCTTAAATATGTTGATGATGATGATGATGATGATGATGATGATGATGATGATG 610  
DB 505 ATGAAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 449  
QY 611 AAAATGAAGATTTAATGCGCAGAAATTTGTTATTTCTTAAAGCGAAGCTCACAGCA 670  
DB 448 AAGAGATGAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 389  
QY 671 CAGTAAATATATCCAGAGATCAAGATGATGATGATGATGATGATGATGATGATG 730  
DB 388 ATGATGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 329  
QY 731 AAGCAAAACGCTTTATCTGGAAGAAATGATGATGATGATGATGATGATGATGATG 790  
DB 328 AAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 269  
QY 791 CAGAGATGGAACGGGAGATTTAAGTTAAATCTTGAAGTGTGCTGATATTTTAG 850

DB 268 ATGAGATGATGATGAGAGAGATGAGATGATGATGAGAGAGATGATGAGAG 209  
QY 851 AAGAGTAAGAGCTCCAAATTAATGCAATTAATGAAAAATCAACAAACACATTTA 910  
DB 208 AAGAGATGAGAGAGAGAGATGATGAGAGATGATGAGAGAGAGATGATGATG 149  
QY 911 CAATTGAGCAACATCAACACCTGTTGAAAAAAGTCAAAATGATTAACCTTAAG 970  
DB 148 AAGAGCAGCAATTAATGATTAATGAGATGAGAGAGATGATGAGAGAGATGAGAG 89  
QY 971 TTGATA 976  
DB 88 AAGATA 83

RESULT 12  
DN711652/c 1416 bp mRNA linear EST 30-MAR-2005  
LOCUS CLJ96-E12 5', mRNA sequence.  
DEFINITION DN711652  
ACCESSION DN711652  
VERSION DN711652.1 GI:62076691  
KEYWORDS EST.  
SOURCE Gasterosteus aculeatus (three spined stickleback).  
ORGANISM Gasterosteus aculeatus  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Actinopterygii; Neopterygii; Teleostei; Euteleostei;  
Acanthomorpha; Acanthopterygii; Percomorpha; Gasterosteiformes;  
Gasterosteidae; Gasterosteus.  
1 (bases 1 to 1416)  
Kingsley, P.M., Petchel, C., Balabhadra, S., Grimwood, J., Dickson, M.,  
Schmutz, J., and Myers, R.M.  
Expressed sequence tags from Gasterosteus aculeatus  
Unpublished (2003)  
Contact: Grimwood, Jane  
Stanford Human Genome Center  
Stanford University School of Medicine  
975 S California Ave, Palo Alto, CA 94304, USA  
Tel: 650 320 5917  
Fax: 650 320 5801  
Email: jane@hgsc.stanford.edu  
Plate: 96  
High quality sequence start: 30  
High quality sequence stop: 303.  
Location/Qualifiers  
1..1416  
/organism="Gasterosteus aculeatus"  
/mol\_type="mRNA"  
/strain="Bilrnfjordur marine sticklebacks, Iceland"  
/db\_xref="taxon:69293"  
/clone="CLJ96-E12"  
/sex="mixed male and female"  
/tissue\_type="whole larva"  
/dev\_stage="21 day old larvae collected at Swarup Stage 30  
/U\_Embryol Exp. Morphol 6: 373-383,1958"  
/lab\_host="DH10B (T1 phage resistant)"  
/clone\_11b="SHGC-CLJ"  
/note="Vector: Express 1; Total and poly A+ RNA was  
isolated from the indicated stickleback tissue, and a cDNA  
library was constructed in the Express 1 plasmid vector by  
Open Biosystems. First strand cDNA synthesis was primed  
with an 54 bp linker primer containing an oligodT sequence  
preceded by a synthetic NotI site (first strand primer:  
5'-GACTAGTCTAGATCGGAGCGCGCCGCTT-3'). Following  
second strand synthesis, cDNAs were made blunt at the end  
corresponding to the original 5' prime end of mRNA, and  
cloned directionally into the NotI and EcoRV sites of  
Express 1. Note that the EcoRV site is typically destroyed  
in the blunt end cloning, leaving a junction of the form  
'xxATC' (where is ATC is the second half of the EcoRV  
site, and xxx is derived from the cDNA sequence). A map of  
the Express 1 vector is available at:

http://www.openbiosystems.com/cdna\_library\_construction\_fa  
q.php#8 The primary library was transformed and amplified  
in DH10B (T1 phage resistant) bacteria. Clones available  
from Open Biosystems:  
http://www.openbiosystems.com/stickleback"

ORIGIN  
Query Match 4.1%; Score 77.6; DB 9; Length 1416;  
Best Local Similarity 39.1%; Pred. No. 4.2e-07;  
Matches 248; Conservative 0; Mismatches 386; Indels 0; Gaps 0;  
QY 485 AATGATCAAGCAACAGATGGTCCATTAATATGCAACAGAAATAGCGGTGTC 544  
DB 1122 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1063  
QY 545 ATATTATCTTAATAATGTGTAGCCATGATGTTTACATGTAAGAAAGTGA 604  
DB 1062 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1003  
QY 605 CTGCTGAATAATGAGATTAATATGCGCGAATTTGTTATTTCTTAAGCGAAGCTC 664  
DB 1002 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 943  
QY 665 CAGCAGAGTAATATATATATCAAGAGTCAAGATGATTAATATATGAGCAACGAGTA 724  
DB 942 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 883  
QY 725 AAGAACAGCAAAACCGCTTATTACTGGAAGAATTGTAAGAAATGATTTCA 784  
DB 882 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 823  
QY 785 CAGAACAGAGAAATGGAACGGAGAAATTAACGTTAAATCTTGAGTTGTTGCTATA 844  
DB 822 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 763  
QY 845 TTTAGAGAGATTAAGCTCCAAATTAATGCAAGATTAATGCAACAAACAC 904  
DB 762 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 703  
QY 905 CATTTCAATTGAGCAAAATCAACACCTGTTGAAAAAAGCTCAAAATGATACCT 964  
DB 702 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 643  
QY 965 CTAAAGTTGATTAACACACACAGCTTAGATGATTAAGATGCGAATGCGGA 1024  
DB 642 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 583  
QY 1025 TTAATATCAATTTCTGTAAATATTCATTTGGGAGTTGCAGACAAAGCGAGCTTA 1084  
DB 582 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 523  
QY 1085 ATAAATACGTCAAATTAATTAATGATTAACA 1118  
DB 522 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 489

RESULT 13  
AZ551092 912 bp DNA linear GSS 14-NOV-2000  
LOCUS ENTJ22TF Entamoeba histolytica Sheared DNA Entamoeba histolytica  
DEFINITION AZ551092  
ACCESSION AZ551092  
VERSION AZ551092.1 GI:11176393  
KEYWORDS GSS.  
SOURCE Entamoeba histolytica  
ORGANISM Entamoeba histolytica  
Eukaryota; Entamoebidae; Entamoeba.  
REFERENCE 1 (bases 1 to 912)  
Loftus, B., Van Aken, S. and Fraser, C.  
Determination of clone end sequences from Entamoeba histolytica  
HMI-IMSS sheared DNA library  
Unpublished (2000)  
JOURNAL COMMENT Contact: Brendan J Loftus

Department of Eukaryotic Genomics  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850, USA  
Tel: 301 838 0208  
Fax: 301 838 3543  
Email: bjoftus@igr.org

Clones are derived from the *Entamoeba histolytica* HMI-IMSS sheared DNA library  
Seq primer: M13-Forward  
Class: shotgun  
High quality sequence start: 17  
High quality sequence stop: 861.  
Location/Qualifiers  
1. 912  
/organism="Entamoeba histolytica"  
/mol\_type="genomic DNA"  
/strain="HMI:IMSS"  
/db\_xref="taxon:5759"  
/note="Vector: pHOSt; Site 1: Bst I; Constructed at The Institute for Genomic Research (TIGR), Rockville, MD. Genomic DNA isolated from broth cultures of *E. histolytica* using a method described by Clark and Diamond (Clark, C.G., and Diamond, L.S. (1993) *Entamoeba histolytica*: a method for isolate identification. *Exp. Parasitol.* 77:450). The DNA was mechanically sheared to give a tight size distribution (~2 kb). The v + i method used for the library construction is described in detail in Smith, H.O. and Venter, J.C. (Making small insert libraries for whole genome shotgun sequencing projects. In *Genome Sequencing: A Practical Approach*, eds. M. Vaudin and B. Barrell, Oxford University Press, 1999)."

## FEATURES

source

Query Match 4.1%; Score 77.2; DB 11; Length 912;  
Best Local Similarity 44.3%; Pred. No. 4.8e-07;  
Matches 316; Conservative 0; Mismatches 398; Indels 0; Gaps 0;

## ORIGIN

518 ATGGAACGAGAAATTTGCGGTTTCTATTTCTTAAATGTGTCGCAATGATG 577  
Db 136 AAGAAGAAATGAAATTTGAAAACTACACAAATGAATTTAGAAAGCGAAGTGAAG 195  
Qy 578 GTAGTTTACATGTGAAAAAGTAGGAACCTGTAATAAGATTAAATGGCGCAAT 637  
Db 196 AATATGATTTGAAGAAGAAAGATGATGATGATGAAGAAGAAAGATGATGAG 255  
Qy 638 TTGTTATTTCTAAAAGCGAAGCTCACGACACAGTAAATATATCCAAAGATCAAAG 697  
Db 256 ATGATGATGATGAAGAAGATGATGATGAAGAAGATGATGAAGAAGATGATGATG 315  
Qy 638 ATGATTTATTCATGACACAGGATTAAGAACCAAGCAACGCTTATTTACTGGAAAA 757  
Db 316 AAGAAGATGATGAAGAAGATGATGAAGAAGATGATGAAGAAGATGATGAAG 375  
Qy 758 GTTATGAATTTGGCAAAATGATTTTCAACGACAGAGATGAACGGGAATTAACAG 817  
Db 376 ACGATGAAGACGACCAATTTGAATTTGAAGAAGATGATGATGAAGAAGAAAGATGATG 435  
Qy 818 TTTAAATCTTGAGGTTGTTGTTATTTTGAAGAAGTAAAGCTCCAAATTAATGCG 877  
Db 436 ATGAAGATGATGAAGACGATGAAGACGAGATTAAGATTAAGATGATGATGATG 495  
Qy 878 AATTATTTGAAATTCAAACAAAAACCCATTACATTTGAACAAACATCAACACCTG 937  
Db 496 AAGAAGAAATGATGATGAAGAAGAAAGATGATGAAGAAGATGATGATGAAGAAG 555  
Qy 938 TTGAAGAAACAGTCAAAATGATACCTTAAAGTTGATTAACCAACACCAAGCTTAAG 997  
Db 556 ATGAAGAAATGATGAAGAAGATGATGAAGAAGATGATGAAGAAGAAAGATGATG 615  
Qy 998 GTAAAGATGTGGCAATGTGGCGAAAAATTTAAATCAATTTCTGTAATATTTCCATTGG 1057  
Db 616 AAGATGATGAAGACATGAAGACGACGAATATGAATTAAGATGATGATGAAGAAG 675

Qy 1058 GGATTCAGACAAAGAGCGCAGCTTAATTAATAGTCAATTTAGTTGATTAAC 1117  
Db 676 AAGATGATGATGAAGAAGAAAGATGATGATGAAGAAGATGATGAAGAAGATGAG 735  
Qy 1118 ATGATGACAGCCCTTAATCTTTGATTAACGTAAGTCTTGTGAGAGATGATCTTAATGCTTATATG 1177  
Db 736 ATGATGATGAAGAAGAAAGATGATGAAGAAGATGATGAAGAAGAAAGATGATGATGAAGAAG 795  
Qy 1178 ATGGGATACAGTATTTGCTCTGAAATTTATCAAGTGAATGAACAGCAATG 1231  
Db 796 ATGAAGAAAGAAAGATGATGAAGAAGATGATGAAGAAGAAAGATGATGATG 849

## RESULT 14

CNS01JRG/c 879 bp DNA linear GSS 12-JUN-2001  
LOCUS  
DEFINITION  
Anopheles gambiae GSS 17 end of clone 14D07 of NotreDamel library from strain PST of Anopheles gambiae (African malaria mosquito), genomic survey sequence.

ACCESSION  
AL147405  
VERSION  
AL147405.1 GI:7005551  
KEYWORDS  
SOURCE  
ORGANISM  
Anopheles gambiae (African malaria mosquito)  
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Preygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Culicidae; Anophelinae; Anopheles.

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
Submitted (16-FEB-2000) Genoscope - Centre National de Sequencage; BP 191 91006 Evry cedex - FRANCE (E-mail: sequefgenoscope.cns.fr  
- Web : www.genoscope.cns.fr)  
2 (bases 1 to 879)  
Roth,C.W., Brey,P.T., Ke,Z., Collins,F.H. and Weissenbach,J.  
Direct Submission  
Submitted (16-FEB-2000) BBMI, Institut Pasteur, 25, rue du Dr.  
Roux, Paris 75015, France  
This clone is from an A. gambiae BAC library provided by F.H. Collins and sequenced by Genoscope in collaboration with the Laboratory of Biochem. and Biol. Molec. of Insects, Institut Pasteur.

COMMENT  
JOURNAL  
Submitted (16-FEB-2000) BBMI, Institut Pasteur, 25, rue du Dr.  
Roux, Paris 75015, France  
This clone is from an A. gambiae BAC library provided by F.H. Collins and sequenced by Genoscope in collaboration with the Laboratory of Biochem. and Biol. Molec. of Insects, Institut Pasteur.

## FEATURES

location/Qualifiers  
1. 879  
/organism="Anopheles gambiae"  
/mol\_type="genomic DNA"  
/strain="PST"  
/db\_xref="taxon:7165"  
/clone="14D07"  
/clone\_lib="Notredamel"  
/note="end : 17"

## ORIGIN

Query Match 4.1%; Score 77; DB 14; Length 879;  
Best Local Similarity 39.9%; Pred. No. 5.3e-07;  
Matches 294; Conservative 47; Mismatches 393; Indels 2; Gaps 2;

385 AAAAAACAAATGTAAAGATGACAGTATACATTAAGAAGAAACAAAGAGGTGTA 444  
Db 792 AAAAAACAAATGTAAAGATGACAGTATACATTAAGAAGAAACAAAGAGGTGTA 733  
Qy 445 GTTGTGCTACGAATATGTGTGCTGCTTCCAGTTTACGA-AAATGATCAAGCAACGA 503  
Db 732 ATAGTAAAAATATGTGAGWADKRAKMAAAAGKTRAAKGMWAAAAAATAATAA 673  
Qy 504 TGGCTCTATTAATATGACAGAAATTAAGCGGTTTCTGATTTATTTCTTAAATATGT 563  
Db 672 AAAAAACAAATGTAAAGATGACAGTATACATTAAGAAGAAACAAAGAGGTGTA 613  
Qy 564 GTTGTGCTACGAATATGTGTGCTGCTTCCAGTTTACGA-AAATGATCAAGCAACGA 622  
Db 612 KGAATAATTTGWWGKTIDAAAAAATTTAGAAAAAAGAGWAGKRGRAAAGAAAGKAK 553





**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.9  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 5, 2006, 18:05:13 ; Search time 826 Seconds  
(without alignments)  
15877.481 Million cell updates/sec

Title: US-10-661-809a-12

Perfect score: 1881

Sequence: 1 atgaagcaatcaaaaaaagt.....gacgragaagaagaatgct 1881

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 524920 seqs, 3486124231 residues

Total number of hits satisfying chosen parameters: 10489840

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database :

N Geneseq\_8: \*  
1: geneseqn1980s: \*  
2: geneseqn1990s: \*  
3: geneseqn2000s: \*  
4: geneseqn2001as: \*  
5: geneseqn2001bs: \*  
6: geneseqn2002as: \*  
7: geneseqn2002bs: \*  
8: geneseqn2003as: \*  
9: geneseqn2003bs: \*  
10: geneseqn2003cs: \*  
11: geneseqn2003ds: \*  
12: geneseqn2004as: \*  
13: geneseqn2004bs: \*  
14: geneseqn2005s: \*  
15: geneseqn2006s: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score  | Query Match | Length | DB ID | Description |
|------------|--------|-------------|--------|-------|-------------|
| 1          | 1881   | 100.0       | 1980   | 10    | ADH84834    |
| 2          | 1881   | 100.0       | 1983   | 2     | AAK20108    |
| 3          | 1881   | 100.0       | 1983   | 6     | ABN98093    |
| 4          | 1881   | 100.0       | 1983   | 6     | ACA88057    |
| 5          | 1881   | 100.0       | 1983   | 5     | ABX61663    |
| 6          | 1881   | 100.0       | 1983   | 14    | ADY38951    |
| 7          | 1881   | 100.0       | 15614  | 2     | AAK12982    |
| 8          | 1881   | 100.0       | 15614  | 6     | ABN98777    |
| 9          | 1859   | 98.8        | 1875   | 12    | ADH84888    |
| 10         | 1859   | 98.8        | 1875   | 14    | ADV16333    |
| 11         | 1687   | 89.7        | 1687   | 2     | AAK20109    |
| 12         | 1687   | 89.7        | 1687   | 6     | ABN98094    |
| 13         | 1687   | 89.7        | 1687   | 6     | ACA88058    |
| 14         | 1687   | 89.7        | 1687   | 8     | ABX61664    |
| 15         | 1687   | 89.7        | 1687   | 14    | ADY38953    |
| 16         | 1014.6 | 53.9        | 1878   | 12    | ADH84894    |
| 17         | 1014.6 | 53.9        | 1878   | 14    | ADV16652    |
| 18         | 1014.6 | 53.9        | 1995   | 10    | ADC93971    |

|    |      |      |        |    |             |                     |
|----|------|------|--------|----|-------------|---------------------|
| 19 | 595  | 31.6 | 2199   | 6  | ABK11591    | Abk11591 E. faecal  |
| 20 | 83.8 | 4.5  | 3399   | 2  | AAT05868    | Aat05868 Chicken 1  |
| 21 | 80   | 4.3  | 1701   | 14 | ADZ72262    | Adz72262 Plasmodu   |
| 22 | 79.8 | 4.2  | 2395   | 15 | AEE62999    | Aee62999 Leucocyto  |
| 23 | 70.4 | 3.7  | 1686   | 2  | AAQ87587    | Aaq87587 DNA encod  |
| 24 | 65   | 3.5  | 627    | 13 | ACN54555    | Acn54555 Cotton an  |
| 25 | 64.8 | 3.4  | 1998   | 8  | AAA70212    | Aaa70212 Plasmodu   |
| 26 | 64.6 | 3.4  | 6609   | 8  | ABZ22899    | Abz22899 Streptococ |
| 27 | 64.2 | 3.4  | 2115   | 6  | ABN67915    | Abn67915 Streptococ |
| 28 | 64.2 | 3.4  | 14006  | 6  | ABJ33958    | Abj33958 Human imm  |
| 29 | 64.2 | 3.4  | 110000 | 6  | ABN71527_13 | Continuaction (14 o |
| 30 | 63.2 | 3.4  | 1509   | 5  | ABD02390    | Abd02390 Virulent   |
| 31 | 63.2 | 3.4  | 1509   | 5  | ABK11585    | Abk11585 S. agalac  |
| 32 | 63   | 3.3  | 822    | 6  | ABO51819    | Abg51819 Oligonuc   |
| 33 | 63   | 3.3  | 822    | 6  | ABQ44852    | Abq44852 Oligonuc   |
| 34 | 63   | 3.3  | 822    | 6  | ABQ44853    | Abq44853 Oligonuc   |
| 35 | 63   | 3.3  | 822    | 6  | ABQ44853    | Abq44853 Oligonuc   |
| 36 | 63   | 3.3  | 3579   | 4  | AAK70099    | Aak70099 Plasmodu   |
| 37 | 61.8 | 3.3  | 9539   | 4  | AAK45347    | Aak45347 Chemical   |
| 38 | 61.8 | 3.3  | 9539   | 6  | ABK28180    | Abk28180 DNA trans  |
| 39 | 61.6 | 3.3  | 516    | 8  | ABK40620    | Abk40620 Bovine ES  |
| 40 | 61   | 3.2  | 700    | 10 | ACD92384    | Adc92384 Human col  |
| 41 | 60.6 | 3.2  | 1065   | 14 | ADZ72252    | Adz72252 Plasmodu   |
| 42 | 60.6 | 3.2  | 1065   | 15 | AEF80820    | Aef80820 Plasmodu   |
| 43 | 60   | 3.2  | 15674  | 6  | ABJ32363    | Abj32363 Human imm  |
| 44 | 60   | 3.2  | 15674  | 6  | ABJ34477    | Abj34477 Human met  |
| 45 | 60   | 3.2  | 15674  | 6  | ABJ70514    | Abj70514 Chemical   |

#### ALIGNMENTS

RESULT 1  
ADH84834  
ID ADH84834 standard; DNA; 1980 BP.  
XX  
XX ADH84834;  
XX  
XX 22-APR-2004 (first entry)  
XX  
XX Enterococcus faecalis polynucleotide #2719.  
XX  
XX Enterococcus faecalis infection; transcription regulatory element;  
XX  
XX antimicrobial; gene; de.  
XX  
XX Enterococcus faecalis.  
XX  
XX US6617156-B1.  
XX  
XX 09-SEP-2003.  
XX  
XX 13-AUG-1998; 98US-00134000.  
XX  
XX 15-AUG-1997; 97US-0055778P.  
XX  
XX (DOUC/) DOUCETTE-STAMM L A.  
XX  
XX (BUSH/) BUSH D.  
XX  
XX Doucette-Stamm LA, Bush D;  
XX  
XX WPI; 2003-895394/82.  
XX  
XX P-PSDB; ADH88239.  
XX  
XX New nucleic acid comprising a sequence encoding an Enterococcus faecalis  
XX  
XX polypeptide, useful for preparing a composition for diagnosing or  
XX  
XX treating E. faecalis infection.  
XX  
XX Disclosure: SEQ ID NO 2719; 193pp; English.  
XX  
XX The invention relates to Enterococcus faecalis polynucleotides and  
XX  
XX polypeptides. The invention also relates to a recombinant expression  
XX  
XX vector comprising a polynucleotide operably linked to a transcription  
XX  
XX regulatory element, a cell comprising a recombinant vector, a method for

CC producing an E. faecalis polypeptide, an isolated nucleic acid comprising  
CC a sequence not given in the specification, a recombinant vector  
CC comprising the nucleic acid and a cell comprising the recombinant vector.  
CC The polynucleotides can be used to detect the presence of E. faecalis in  
CC a sample. The sequences are useful for preparing a composition for  
CC diagnosing or treating Enterococcus faecalis infection. This sequence  
CC represents an E. faecalis polynucleotide of the invention.

XX Sequence 1980 BP; 728 A; 330 C; 425 G; 497 T; 0 U; 0 Other;

Query Match 100.0%; Score 1881; DB 10; Length 1980;

Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1881; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 ATGAAGCAATTAAAAAGTTTGGTACACCGTTAGTACCTTGTACTAATTTTGGCAGCTT 60
DB 97 ATGAAGCAATTAAAAAGTTTGGTACACCGTTAGTACCTTGTACTAATTTTGGCAGCTT 156
QY 61 TTCAAGGTATTTGGGCAACAACCTGATTTGCGAAGAAAATGGGGAGAGCGCAG 120
DB 157 TTCAAGGTATTTGGGCAACAACCTGATTTGCGAAGAAAATGGGGAGAGCGCAG 216
QY 121 CTGCGATTCAAAAAGAAAAATGACGATTTACAGATCCGCTTAATTCAAAATAGCGGG 180
DB 217 CTGCGATTCAAAAAGAAAAATGACGATTTACAGATCCGCTTAATTCAAAATAGCGGG 276
QY 181 AAGAAATGAGCGAGTTTGTATTAATTCAGAGCTGGCGATGTGACGTTTAGTATTAT 240
DB 277 AAGAAATGAGCGAGTTTGTATTAATTCAGAGCTGGCGATGTGACGTTTAGTATTAT 336
QY 241 AACGTGACGAACGAATTTTACGAGCAACGAGCGGCGAAGCGTTGATGACAGTANA 300
DB 337 AACGTGACGAACGAATTTTACGAGCAACGAGCGGCGAAGCGTTGATGACAGTANA 396
QY 301 CAAGCTGTCCAAAGTTTAACTCTCGGGAAAACCTGTTCTCAAGAAACCAACGATGCAAT 360
DB 397 CAAGCTGTCCAAAGTTTAACTCTCGGGAAAACCTGTTCTCAAGAAACCAACGATGCAAT 456
QY 361 GGGAAATGTCACCTGTTCACTTACCTTAAAAACAAAATGTTAAGATGCACTGTATACAT 420
DB 457 GGGAAATGTCACCTGTTCACTTACCTTAAAAACAAAATGTTAAGATGCACTGTATACAT 516
QY 421 AAGAAAGAACCAAAAAGAGGGTGTAGTGTGCTACGATATGTTGGTGGCGTTCCAGTT 480
DB 517 AAGAAAGAACCAAAAAGAGGGTGTAGTGTGCTACGATATGTTGGTGGCGTTCCAGTT 576
QY 481 TAGCAATGATCAAGCAAAACAGATGGTCTCTATAATATGAAACAGAAAGATTAAGCGTT 540
DB 577 TAGCAATGATCAAGCAAAACAGATGGTCTCTATAATATGAAACAGAAAGATTAAGCGTT 636
QY 541 GTTCAATTTTATCTTAAAAATGTGGTAGCCCAATGATGTATGTTTACATGTGAAAAAGTA 600
DB 637 GTTCAATTTTATCTTAAAAATGTGGTAGCCCAATGATGTATGTTTACATGTGAAAAAGTA 696
QY 601 GGAACGTGCGAAAAAGAGGATTTAAATGGCGAGAAATTTGTTATTTCTAAAAAGCGAAGG 660
DB 697 GGAACGTGCGAAAAAGAGGATTTAAATGGCGAGAAATTTGTTATTTCTAAAAAGCGAAGG 756
QY 661 TCACCAAGGACAGTAATAATATATCCAAAGAGTCAAAAGATGATATATATCATGACAAAG 720
DB 757 TCACCAAGGACAGTAATAATATATCCAAAGAGTCAAAAGATGATATATATCATGACAAAG 816
QY 721 GATTAAGAACAAAGCAAAACGCTTTATTTACTGGGAAAAGTTATGAAATTTGGCGAAAATGAT 780
DB 817 GATTAAGAACAAAGCAAAACGCTTTATTTACTGGGAAAAGTTATGAAATTTGGCGAAAATGAT 876
QY 781 TTCACAGAGACAGAAATGAAACGGGAGATTAAACGTTAAAAATTTGAGGTTGGTTGG 840
DB 877 TTCACAGAGACAGAAATGAAACGGGAGATTAAACGTTAAAAATTTGAGGTTGGTTGG 936
QY 841 TATATTTTGAAGAAAGTAAAGCTCCAAATTAATGCAAGATTATGAAAAATCAAAACAAA 900
DB 937 TATATTTTGAAGAAAGTAAAGCTCCAAATTAATGCAAGATTATGAAAAATCAAAACAAA 996
```

```
QY 901 ACACATTTAACAATTGAAGCAAAACAAATCAAAACACCTGTTGAAAAAAGAGTCAAAAATGAT 960
DB 997 ACACATTTAACAATTGAAGCAAAACAAATCAAAACACCTGTTGAAAAAAGAGTCAAAAATGAT 1056
QY 961 ACCTTAAAGTTGATTAACCAACACCAAGCTTGAATGGTAAAGATGTGGCAATTGGCGAA 1020
DB 1057 ACCTTAAAGTTGATTAACCAACACCAAGCTTGAATGGTAAAGATGTGGCAATTGGCGAA 1116
QY 1021 AAAATTAATATCAAAATTTCTGTAAATATATTCATTTGGGATTTGCGAGCAAAAGCGAC 1080
DB 1117 AAAATTAATATCAAAATTTCTGTAAATATATTCATTTGGGATTTGCGAGCAAAAGCGAC 1176
QY 1081 GCTAATAATATGTCAAATTCATTTAGTGAATAACATGATGACGCTTAATCTTTGAT 1140
DB 1177 GCTAATAATATGTCAAATTCATTTAGTGAATAACATGATGACGCTTAATCTTTGAT 1236
QY 1141 AACGTACTTCTGAGAGATATGCTTATGCTTATATATGATGGGATATCAGTATGCTCCT 1200
DB 1237 AACGTACTTCTGAGAGATATGCTTATGCTTATATGATGGGATATCAGTATGCTCCT 1296
QY 1201 GAAATTTATCAAGTGACTGGAACACAAATGAGCTTCACTGTGCGCTTAATTCACGATAT 1260
DB 1297 GAAATTTATCAAGTGACTGGAACACAAATGAGCTTCACTGTGCGCTTAATTCACGATAT 1356
QY 1261 ATTCTAGGCTTAACACAGGCGGCACTAAATTCGTTTACTTATGCAATTTAATGAA 1320
DB 1357 ATTCTAGGCTTAACACAGGCGGCACTAAATTCGTTTACTTATGCAATTTAATGAA 1416
QY 1321 AAGCAGATCTTACGAAAGGCTTTAAAAATGAGGGAATGTTGATPACGGTCATACCGAC 1380
DB 1417 AAGCAGATCTTACGAAAGGCTTTAAAAATGAGGGAATGTTGATPACGGTCATACCGAC 1476
QY 1381 GACCAAAACCAACCACTGTTGAATGTGACAGGTGGGAAAACGTTTCAATTAATCCAT 1440
DB 1477 GACCAAAACCAACCACTGTTGAATGTGACAGGTGGGAAAACGTTTCAATTAATCCAT 1536
QY 1441 GGGCATGTGACAGCACAACAGCTTGGGAGGAGCTTCTTGTGCTGCTGATCAAAAAC 1500
DB 1537 GGGCATGTGACAGCACAACAGCTTGGGAGGAGCTTCTTGTGCTGCTGATCAAAAAC 1596
QY 1501 AGCGACACGCAAAATTTATGAAAAATGATGAACAAACGAAAGACGAACTTGGGTGAAA 1560
DB 1597 AGCGACACGCAAAATTTATGAAAAATGATGAACAAACGAAAGACGAACTTGGGTGAAA 1656
QY 1561 ACAAAGCTGAAGCAACTTCTTACACAAACAGGCTGATGATGTTGATATCAACAGG 1620
DB 1657 ACAAAGCTGAAGCAACTTCTTACACAAACAGGCTGATGATGTTGATATCAACAGG 1716
QY 1621 CTTAAATACGGTACCTAATTTATGAAAGAACTGAGCTCTGATGATATGCTTGTGA 1680
DB 1717 CTTAAATACGGTACCTAATTTATGAAAGAACTGAGCTCTGATGATATGCTTGTGA 1776
QY 1681 ACAAAATCGAATGGAATTTGTGTGATGATGAACATGATATGGAACAAAGAAACCTAGTT 1740
DB 1777 ACAAAATCGAATGGAATTTGTGTGATGATGAACATGATATGGAACAAAGAAACCTAGTT 1836
QY 1741 TCACAGAAAAAGTACCAAAACAAACAAAGGTACCTTAACCTTCAACAGGTGGCAAGGA 1800
DB 1837 TCACAGAAAAAGTACCAAAACAAACAAAGGTACCTTAACCTTCAACAGGTGGCAAGGA 1896
QY 1801 ATCTACGTTTACTTGAAGAGTGGCGAGCTTGTACTTAATTTGCGAGAGTCACTTTGCT 1860
DB 1897 ATCTACGTTTACTTGAAGAGTGGCGAGCTTGTACTTAATTTGCGAGAGTCACTTTGCT 1956
QY 1861 AGACGTAGAAAAAGAAATGCT 1881
DB 1957 AGACGTAGAAAAAGAAATGCT 1977
```

RESULT 2  
AAAX20108  
ID AAAX20108 standard; DNA; 1983 BP.

XX AAX20108;  
AC 20-APR-1999 (first entry)  
XX DE Enterococcus faecalis gene EF058.  
XX DE Enterococcus faecalis  
XX DE Enterococcus faecalis; infection; vaccine; immune response; diagnosis;  
XX DE detection; attenuation; antigenic; ss.  
XX DE Enterococcus faecalis.  
XX OS WO9850554-A2.  
XX PD 12-NOV-1998.  
XX PF 04-MAY-1998; 98WO-US008959.  
XX PR 06-MAY-1997; 97US-0044031P.  
XX PR 16-MAY-1997; 97US-004655P.  
XX PR 14-NOV-1997; 97US-006609P.  
XX PA (HUMA-) HUMAN GENOME SCI INC.  
XX PI Kunsch CA, Choi GH, Bailey C, Hromockyj A;  
XX WI: 1999-070095/06.  
XX DR P-PSDB; AAY00118.  
XX PS New isolated Enterococcus faecalis polynucleotides - used to develop  
XX PT products for the detection of Enterococcus and for use in vaccines for  
XX PT prevention or attenuation of Enterococcus infection.  
XX PS Claim 1; Page 133; 301pp; English.  
XX PS The present sequence represents a gene isolated from Enterococcus  
XX CC faecalis. The present invention describes genes, proteins and antigenic  
XX CC polypeptides isolated from B. faecalis. The proteins can be used in  
XX CC vaccines for preventing or attenuating an infection caused by a member of  
XX CC the Enterococcus genus in an animal. They can also be used for detecting  
XX CC Enterococcus antibodies in a sample. The nucleotide sequences can be used  
XX CC for detecting Enterococcus nucleic acids. Products from the present  
XX CC invention can also be used for screening compounds to identify agonists  
XX CC and antagonists of E. faecalis protein activity  
XX CC  
XX SQ Sequence 1983 BP; 729 A; 330 C; 426 G; 498 T; 0 U; 0 Other;  
Query Match 100.0%; Score 1881; DB 2; Length 1983;  
Best Local Similarity 100.0%; Pred. No. 0; Mismatches 0; Indels 0; Gaps 0;  
Matches 1881; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ATGAAGCAATTAATAAGTTGGTACACCGTTAGTACTTGTCTAATTTGGCACTT 60  
DB 100 ATGAAGCAATTAATAAGTTGGTACACCGTTAGTACTTGTCTAATTTGGCACTT 159  
QY 61 TTCACAACTGTTATGGGACAACTGCTTTGGGAGAGAAATGGGGAGAGCCCAAG 120  
DB 160 TTCACAACTGTTATGGGACAACTGCTTTGGGAGAGAAATGGGGAGAGCCCAAG 219  
QY 121 CTCGTGATTCACAAAATAAGTACGATTTACAGATCCGCTTAATCAAAATAGCGG 180  
DB 220 CTCGTGATTCACAAAATAAGTACGATTTACAGATCCGCTTAATCAAAATAGCGG 279  
QY 181 AAAGAATGACGAGTTGATTAATATCAAGGACTGGAGATGTGACGTTTATAT 240  
DB 280 AAAGAATGACGAGTTGATTAATATCAAGGACTGGAGATGTGACGTTTATAT 339  
QY 241 AACGGAAGCAATTTTATGAGAGAGAGGCGGCAAGGCTTGATGACACTAA 300  
DB 340 AACGGAAGCAATTTTATGAGAGAGAGGCGGCAAGGCTTGATGACACTAA 399  
QY 301 CAAGCTGCCAAAGTTTAACTCTGGGAAACCTGTTGCTCAAGGAACGAGTCAAT 360  
|||||

DB 400 CAAGCTGCCAAAGTTTAACTCTGGGAAACCTGTTGCTCAAGGAACGAGTCAAT 459  
QY 361 GGAATGTCACCTGTTCACTTACCTTAAATAAATAATGTAAGATGCACTGATACATT 420  
DB 460 GGAATGTCACCTGTTCACTTACCTTAAATAAATAATGTAAGATGCACTGATACATT 519  
QY 421 AAAGAAGCAACCAAAAGAGGTAGTCTGCTGACGAATATGAGTGGCGTTCCAGT 480  
DB 520 AAAGAAGCAACCAAAAGAGGTAGTCTGCTGACGAATATGAGTGGCGTTCCAGT 579  
QY 481 TACGAATGATCAAGCAACAGATGTTCTTAATAATGAAACAGAAATATGACGTT 540  
DB 580 TACGAATGATCAAGCAACAGATGTTCTTAATAATGAAACAGAAATATGACGTT 639  
QY 541 GTTCATATTTATCTTAAATAATGCTGACGCAATGATGTTATCATGTGAAAAAGTA 600  
DB 640 GTTCATATTTATCTTAAATAATGCTGACGCAATGATGTTATCATGTGAAAAAGTA 699  
QY 601 GGAATGCTGAAAAATGAAAGATTAAATGCGGAGAAATTTGTTAATCTTAAAGCGAGC 660  
DB 700 GGAATGCTGAAAAATGAAAGATTAAATGCGGAGAAATTTGTTAATCTTAAAGCGAGC 759  
QY 661 TCACGAGCAGACGTAATAATATATCCAGAGATCAAGATGATTAATATCATGACAAAG 720  
DB 760 TCACGAGCAGACGTAATAATATATCCAGAGATCAAGATGATTAATATCATGACAAAG 819  
QY 721 GATTAAGAACAGCAAAACGCTTTATTAATGCGGAAATGTTATGAATGGCGAAATGAT 780  
DB 820 GATTAAGAACAGCAAAACGCTTTATTAATGCGGAAATGTTATGAATGGCGAAATGAT 879  
QY 761 TTCACAGAGCAGAGAAATGGAACCGGAGAAATTAACGTTAAATCTTGAGGTGTTG 840  
DB 880 TTCACAGAGCAGAGAAATGGAACCGGAGAAATTAACGTTAAATCTTGAGGTGTTG 939  
QY 841 TATATTTTGAAGAAAGTAAAGCTCCAAATATGACGATTAATTAATCAACAA 900  
DB 940 TATATTTTGAAGAAAGTAAAGCTCCAAATATGACGATTAATTAATCAACAA 999  
QY 901 ACACATTTTACAAATGAGCAACCAATCAACACTGTTGAAAAACGTCAAAAATGAT 960  
DB 1000 ACACATTTTACAAATGAGCAACCAATCAACACTGTTGAAAAACGTCAAAAATGAT 1059  
QY 961 ACCTTAAGTTGAATTAACCAACCAAGCTTGAATGTAAGATGAGCAATGGCGAA 1020  
DB 1060 ACCTTAAGTTGAATTAACCAACCAAGCTTGAATGTAAGATGAGCAATGGCGAA 1119  
QY 1021 AAATTAATATCAAAATTTCTGTAATATTCATTTGGGAAATGCGAACAAGAGCGAC 1080  
DB 1120 AAATTAATATCAAAATTTCTGTAATATTCATTTGGGAAATGCGAACAAGAGCGAC 1179  
QY 1081 GCTAATTAATATCAAAATTTCTGTAATATTCATTTGGGAAATGCGAACAAGAGCGAC 1140  
DB 1180 GCTAATTAATATCAAAATTTCTGTAATATTCATTTGGGAAATGCGAACAAGAGCGAC 1239  
QY 1141 AACGTAATTTGAGAGATATGCTTATGCTTAATGATGAGGAGATACAGTATGCTCCT 1200  
DB 1240 AACGTAATTTGAGAGATATGCTTATGCTTAATGATGAGGAGATACAGTATGCTCCT 1299  
QY 1201 GAAATTAATCAAGTACGTAACAAGCAATGAGCTTCACTGTCGCGCTTAATCAGCGAT 1260  
DB 1300 GAAATTAATCAAGTACGTAACAAGCAATGAGCTTCACTGTCGCGCTTAATCAGCGAT 1359  
QY 1261 ATTCTACGCTAACCAAGGCGGCACTAAATTCGTTTACTTTATGATTAATGAA 1320  
DB 1360 ATTCTACGCTAACCAAGGCGGCACTAAATTCGTTTACTTTATGATTAATGAA 1419  
QY 1321 AAAGCAATCTCAAGAAAGGCTTTAAATAATGAGGCAATGTTGATPAAGGATACGCG 1380  
DB 1420 AAAGCAATCTCAAGAAAGGCTTTAAATAATGAGGCAATGTTGATPAAGGATACGCG 1479  
QY 1381 GACCAACACCAACCAATGTTGAAGTTGTAAGGTGGAAACGTTTATTAAGTCAT 1440  
DB 1480 GACCAACACCAACCAATGTTGAAGTTGTAAGGTGGAAACGTTTATTAAGTCAT 1539  
|||||

QY 1441 GGGCATGTGACAGCAGCAAGAGCTTGGCGGAGCTTCCTTGTGTCGATGCAAAAC 1500  
DB 1540 GGGCATGTGACAGCAGCAAGAGCTTGGCGGAGCTTCCTTGTGTCGATGCAAAAC 1599  
QY 1501 AGCGACAGCAAGCAATTTATTTGAAATCGATGAAACAGAAAGCAAGCACTTGGGTGAAA 1560  
DB 1600 AGCGACAGCAAGCAATTTATTTGAAATCGATGAAACAGAAAGCAAGCACTTGGGTGAAA 1659  
QY 1561 ACAAAGCTGAAAGCACTACTTTTACACACAGCGCTGATGATTAGTTGATATCAAGGG 1620  
DB 1660 ACAAAGCTGAAAGCACTACTTTTACACACAGCGCTGATGATTAGTTGATATCAAGGG 1719  
QY 1621 CTAAATACGGTACTATTTATTTAGAAAGAACTGAGCTCCGATGATTATGCTTGTGA 1680  
DB 1720 CTAAATACGGTACTATTTATTTAGAAAGAACTGAGCTCCGATGATTATGCTTGTGA 1779  
QY 1681 ACAAATCGGATTTGATTTGTGTGTCATGATGAAATGAGCAACAGAAAGCACTAGTT 1740  
DB 1780 ACAAATCGGATTTGATTTGTGTGTCATGATGAAATGAGCAACAGAAAGCACTAGTT 1839  
QY 1741 TCACCAAGAAAAGTACCAACAAACAGAGTACCTTACCTTCAACAGGTGGCAAGGA 1800  
DB 1840 TCACCAAGAAAAGTACCAACAAACAGAGTACCTTACCTTCAACAGGTGGCAAGGA 1899  
QY 1801 ATCTACGTTTACTTGAAGTGGCGCAGCTTGTCTACTTATTTGCAAGAGTCTACTTGTCT 1860  
DB 1900 ATCTACGTTTACTTGAAGTGGCGCAGCTTGTCTACTTATTTGCAAGAGTCTACTTGTCT 1959  
QY 1861 AGACGTAGAAAAGAAAATGCT 1881  
DB 1960 AGACGTAGAAAAGAAAATGCT 1980

## RESULT 3

ABN98093  
ID ABN98093 standard; DNA; 1983 BP.

XX AC ABN98093;

XX DT 05-AUG-2002 (first entry)

XX DE E faecalis EF058 gene.

XX KW Enterococcus; vaccine; gastrointestinal disease; diagnosis; antibiotic;

XX KM gene; ds.

XX OS Enterococcus faecalis.

XX PN US2002045737-A1.

XX PD 18-APR-2002.

XX PF 04-MAY-1998; 98US-00071035.

XX PR 04-MAY-1998; 98US-00071035.

XX PA (HUMA-) HUMAN GENOME SCI INC.

XX PI Choi GH, Bailey C, Hromocky J, Kunsch CA;

XX DR WPI: 2002-425450/45.

XX DR P-PSDB; ABP43337.

XX PT New genes and polypeptides from Enterococcus faecalis, useful as vaccines

XX PT for preventing, treating or attenuating an infection caused by a member

XX PT of the Enterococcus genus in an animal, particularly E. faecalis.

XX PS Claim 1; Page 92; 255pp; English.

XX CC The present invention provides the protein and coding sequences of a

XX CC number of polypeptides from Enterococcus faecalis. The proteins can be

CC member of the Enterococcus genus in an animal, particularly E. faecalis.  
CC The polynucleotide is also useful for preventing or treating E. faecalis  
CC infection. The present sequence is a coding sequence of the invention  
XX  
SQ Sequence 1983 BP; 729 A; 330 C; 426 G; 498 T; 0 U; 0 Other;

Query Match 100.0%; Score 1881; DB 6; Length 1983;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1881; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAAGCAATTAATAAAAGTTTGGTACACCGTTAGTCTTGTATCTAATTTTGGCACTT 60  
DB 100 ATGAAGCAATTAATAAAAGTTTGGTACACCGTTAGTCTTGTATCTAATTTTGGCACTT 159  
QY 61 TTCACAGTGTATTTAGGACAAACATGCACTTGTGAGAAAGAAATGGGAGAGCGCACAG 120  
DB 160 TTCACAGTGTATTTAGGACAAACATGCACTTGTGAGAAAGAAATGGGAGAGCGCACAG 219  
QY 121 CTCGTGATTCACAAAAGAAATGACGAGATTTCACGATCCGCTTATTCAAAATAGCGGG 180  
DB 220 CTCGTGATTCACAAAAGAAATGACGAGATTTCACGATCCGCTTATTCAAAATAGCGGG 279  
QY 181 AAAGAAATGACGAGATTGTGATTAATATCAAGGACTGGCAAGATGACCTTTATGTTAT 240  
DB 280 AAAGAAATGACGAGATTGTGATTAATATCAAGGACTGGCAAGATGACCTTTATGTTAT 339  
QY 241 AACGTGAGAAAGCAATTTTATGAGCAAGAGCGGAGCGCAAGCGTTGATGACCTAA 300  
DB 340 AACGTGAGAAAGCAATTTTATGAGCAAGAGCGGAGCGCAAGCGTTGATGACCTAA 399  
QY 301 CAAGCTGTCAAAAGTTTAACTCTGGGAAACCTGTGTCTCAAGGAAACCAAGCATGCAAT 360  
DB 400 CAAGCTGTCAAAAGTTTAACTCTGGGAAACCTGTGTCTCAAGGAAACCAAGCATGCAAT 459  
QY 361 GGGAAATGCACTGTTCACTTACCTTAAACCAAAATGTTAAAGATGATATCCATT 420  
DB 460 GGGAAATGCACTGTTCACTTACCTTAAACCAAAATGTTAAAGATGATATCCATT 519  
QY 421 AAAGAAAGCAAAAGAGGTGTAGTGTCTGTACGAATATGATGATGATGATGATGATGAT 480  
DB 520 AAAGAAAGCAAAAGAGGTGTAGTGTCTGTACGAATATGATGATGATGATGATGATGAT 579  
QY 541 GTTCATATTTATCTTAAATAATGTGTAGCAATGATGATGATGATGATGATGATGATGAT 600  
DB 580 GTTCATATTTATCTTAAATAATGTGTAGCAATGATGATGATGATGATGATGATGATGAT 699  
QY 601 GGAATGCTGAAAGATGAAAGATTAATGAGCGCAGAAATTTGTTATTTCTTAAAGCAAGGC 660  
DB 700 GGAATGCTGAAAGATGAAAGATTAATGAGCGCAGAAATTTGTTATTTCTTAAAGCAAGGC 759  
QY 661 TCACCAAGCAGAGTAAATATATCCAGAGTCAAGATGATTAATATATATATATATATATAT 720  
DB 760 TCACCAAGCAGAGTAAATATATCCAGAGTCAAGATGATTAATATATATATATATATAT 819  
QY 721 GATTAAGAACAGCAAAAGCGCTTATTTACCTGGGAAAAGTTATGAATTTGGCAAAATGAT 780  
DB 820 GATTAAGAACAGCAAAAGCGCTTATTTACCTGGGAAAAGTTATGAATTTGGCAAAATGAT 879  
QY 781 TTCACAGAGCAGAGATGAAACCGGAGAAATTAACGTTAATAAATCTTGAGGTGCTCG 840  
DB 880 TTCACAGAGCAGAGATGAAACCGGAGAAATTAACGTTAATAAATCTTGAGGTGCTCG 939  
QY 841 TATATTTAGAGAGTAAAGTCAATATATATATATATATATATATATATATATATATATAT 900  
DB 940 TATATTTAGAGAGTAAAGTCAATATATATATATATATATATATATATATATATATATAT 999  
QY 901 ACACCATTTACATTTGAAGCAACATCAACACTGTTGAAAAACGTCAAAATATGAT 960  
DB 1000 ACACCATTTACATTTGAAGCAACATCAACACTGTTGAAAAACGTCAAAATATGAT 1059

```

QY 961 ACCTCTAAAGTTGATTAACCAACCAAGCTAGATGTAAGATGTCGCAATGGCGAA 1020
DB 1060 ACCTCTAAAGTTGATTAACCAACCAAGCTAGATGTAAGATGTCGCAATGGCGAA 1119
QY 1021 AAAATTAATTAATCAAAATTTCTGTAAATATTCATTTGGGAGATGGCAGCAAGAGCGCAG 1080
DB 1120 AAAATTAATTAATCAAAATTTCTGTAAATATTCATTTGGGAGATGGCAGCAAGAGCGCAG 1179
QY 1081 GGTATTAATTAATCAAAATTTCTGTAAATATTCATTTGGGAGATGGCAGCAAGAGCGCAG 1140
DB 1180 GGTATTAATTAATCAAAATTTCTGTAAATATTCATTTGGGAGATGGCAGCAAGAGCGCAG 1239
QY 1141 AACGTACTCTGAGAGATGCTTATGCTTATATGATGGGAGATGCAAGATGCTGCT 1200
DB 1240 AACGTACTCTGAGAGATGCTTATGCTTATATGATGGGAGATGCAAGATGCTGCT 1299
QY 1201 GAAATTAATTAATCAAAATTTCTGTAAATATTCATTTGGGAGATGGCAGCAAGAGCGCAG 1260
DB 1300 GAAATTAATTAATCAAAATTTCTGTAAATATTCATTTGGGAGATGGCAGCAAGAGCGCAG 1359
QY 1261 ATTCTAGCTTAACCAAGAGCTGCTTAAATTTCTGTAAATATTCATTTGGGAGATGGCAGCAAGAGCGCAG 1320
DB 1360 ATTCTAGCTTAACCAAGAGCTGCTTAAATTTCTGTAAATATTCATTTGGGAGATGGCAGCAAGAGCGCAG 1419
QY 1321 AAAGCAGATCTTACGAAAGGCTTTAAATTAAGGCGAATGTTGATTAAGCTATCCGAC 1380
DB 1420 AAAGCAGATCTTACGAAAGGCTTTAAATTAAGGCGAATGTTGATTAAGCTATCCGAC 1479
QY 1381 GACCAAAACCAACCAAGCTGTTGAAGTTGACAGGTGGGAAAGCTTCAATTAAGTCGAT 1440
DB 1480 GACCAAAACCAACCAAGCTGTTGAAGTTGACAGGTGGGAAAGCTTCAATTAAGTCGAT 1539
QY 1441 GGGCATGTGACAGCGACACCAACCTTTGGGAGGCTTCTTTGTGTGCTCCGTATCAAAAC 1500
DB 1540 GGGCATGTGACAGCGACACCAACCTTTGGGAGGCTTCTTTGTGTGCTCCGTATCAAAAC 1599
QY 1501 AGCGACACGCAAAATTTATTTGAAATTCGATGAAACCAAGAAAGCAAGCTTGGGAGAA 1560
DB 1600 AGCGACACGCAAAATTTATTTGAAATTCGATGAAACCAAGAAAGCAAGCTTGGGAGAA 1659
QY 1561 ACAAAAGCTGAAGCAACTTCTTTTCAACCAAGCGGTGATGATTAAGTATCAAGAG 1620
DB 1660 ACAAAAGCTGAAGCAACTTCTTTTCAACCAAGCGGTGATGATTAAGTATCAAGAG 1719
QY 1621 CTTAAATACGCTTATTTATTAAGAAACTGTAGCTCCGTAGATTAATGTCTTTGTTA 1680
DB 1720 CTTAAATACGCTTATTTATTAAGAAACTGTAGCTCCGTAGATTAATGTCTTTGTTA 1779
QY 1681 ACAATTCGATTTGATTTGTGTCATATGATGCAATGATGCAACCAAGAAACCTAGTT 1740
DB 1780 ACAATTCGATTTGATTTGTGTCATATGATGCAATGATGCAACCAAGAAACCTAGTT 1839
QY 1741 TCACCAAGAAAGTACCAAAACCAAGCAAGAGTACCTTCAACAGGTGGCAAGGA 1800
DB 1840 TCACCAAGAAAGTACCAAAACCAAGCAAGAGTACCTTCAACAGGTGGCAAGGA 1899
QY 1801 ATCTACGTTTACTTGAAGTGGCGCAGCTTGTCTACTTATTTAGCAGAGTCTACTTTGCT 1860
DB 1900 ATCTACGTTTACTTGAAGTGGCGCAGCTTGTCTACTTATTTAGCAGAGTCTACTTTGCT 1959
QY 1861 AGACGTAGAAAAGAAATGCT 1881
DB 1960 AGACGTAGAAAAGAAATGCT 1980

```

```

RESULT 4
ID ACA88057
XX ACA88057 standard; DNA; 1983 BP.
AC ACA88057;
XX
DT 07-JUL-2003 (first entry)

```

```

XX DE E. faecalis novel gene #109.
XX KW Gene; da; endocarditis; bacteraemia; urinary tract infection; UTI;
XX KW intraabdominal infection; soft tissue infection; neonatal sepsis;
XX KW vaccine.
XX OS Enterococcus faecalis.
XX PN US2003017495-A1.
XX PD 23-JAN-2003.
XX PF 29-JUL-2002; 2002US-00206576.
XX PR 06-MAY-1997; 97US-0044031P.
XX PR 16-MAY-1997; 97US-0046655P.
XX PR 14-NOV-1997; 97US-0066009P.
XX PR 04-MAY-1998; 98US-00071035.
XX PA (HUMA-) HUMAN GENOME SCI INC.
XX PI Choi GH, Bailey C, Hromockyj A, Kunsch CA;
XX PT MPI; 2003-416890/39.
XX PT P-PSDB; ABU88365.
XX PT New nucleic acid molecules and polypeptides from Enterococcus faecalis,
XX PT useful as vaccines for preventing or attenuating an enterococcal
XX PT infection in an animal, or for identifying Enterococcus faecalis in
XX PT biological samples.
XX PS Claim 1; Page; 40pp; English.
XX CC The invention relates to a new isolated nucleic acid molecule comprising
XX CC a polynucleotide isolated from Enterococcus faecalis appearing as
XX CC ACA87949-ACA88196 (or sequences complementary to them or 95% identical to
XX CC them). Also included are the proteins encoded by the above nucleic acids,
XX CC making a recombinant vector (comprising inserting the isolated nucleic
XX CC acid molecule cited above into a vector), a host cell comprising the
XX CC vector (used to produce the protein), an isolated antibody specific for
XX CC the polypeptides, a hybridoma that produces the antibody, an isolated
XX CC polypeptide antigen comprising an amino acid sequence of an Enterococcus
XX CC faecalis epitope listed in the specification, a vaccine comprising one or
XX CC more E. faecalis polypeptides (and a pharmaceutical diluent carrier or
XX CC excipient) where the polypeptide elicits protective antibodies in an
XX CC animal to a member of the genus Enterococcus; preventing or attenuating
XX CC an infection caused by a member of the genus Enterococcus in an animal
XX CC comprising administering to the animal the polypeptide and detecting
XX CC Enterococcus nucleic acids in a biological sample. The E. faecalis
XX CC nucleic acid molecules and polypeptides are useful as vaccines for
XX CC preventing or attenuating an enterococcal infection in an animal (e.g.
XX CC endocarditis, bacteraemia, urinary tract infection (UTI), intraabdominal
XX CC infection, soft tissue infection and neonatal sepsis). The polypeptides
XX CC are also useful for detecting Enterococcus aureus in immunoassays, as
XX CC epitope tags, as molecular weight markers, or for generating antibodies
XX CC that specifically bind E. faecalis polypeptides. The nucleic acid
XX CC molecules are also useful as probes for gene mapping, or for identifying
XX CC E. faecalis in biological samples. The kit and methods are useful for
XX CC detecting Enterococcus antibodies or nucleic acid molecules in a
XX CC biological sample. The present sequence is a novel E. faecalis nucleic
XX CC acid of the invention. Note: The sequence data for this patent did not
XX CC form part of the printed specification, but was obtained in electronic
XX CC format directly from the USPTO at
XX CC seqdata.uspto.gov/sequence.html?docID=20030017495
XX SQ Sequence 1983 BP; 729 A; 330 C; 426 G; 498 T; 0 U; 0 Other;
XX
XX Query Match 100.0%; Score 1881; DB 8; Length 1983;
XX Beest Local Similarity 100.0%; Pred. No. 0;
XX Matches 1881; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGAAGCAATTAAAGTTGGTACACCGTTAGTACCTTGTACTAATTTGGCACCTT 60

```

Db 100 ATGAAGCAATTTAAAAAGTTTGGTACACCGTTAGTACCTTGTACTAATTTTGGCACTT 159  
Qy 61 TTCAACAAGTATTTAGGGAACAACAATGCAATTTGAGAGAAATGGGAGAGCGCAAG 120  
Db 160 TTCAACAAGTATTTAGGGAACAACAATGCAATTTGAGAGAAATGGGAGAGCGCAAG 219  
Qy 121 CTCGGATTTCAAAAAAGAAATGACGGATTTTACAGATCCGGCTTATTTCAAAATGCGGG 180  
Db 220 CTCGGATTTCAAAAAAGAAATGACGGATTTTACAGATCCGGCTTATTTCAAAATGCGGG 279  
Qy 181 AAAAAGATGAGCGAGTTGTATTAATATCAAGGACTGGCAGATGTGACGTTTATTTAT 240  
Db 280 AAAAAGATGAGCGAGTTGTATTAATATCAAGGACTGGCAGATGTGACGTTTATTTAT 339  
Qy 241 AACGTGACAAACGAATTTTACGAGCAACGAGCGGCAAGCGCTTATGATGCACTAAA 300  
Db 340 AACGTGACAAACGAATTTTACGAGCAACGAGCGGCAAGCGCTTATGATGCACTAAA 399  
Qy 301 CAAGCTGTCCAAAGTTTAACTCTCGGGAAACCTGTTCCTCAAGAAACCAAGATGCAAT 360  
Db 400 CAAGCTGTCCAAAGTTTAACTCTCGGGAAACCTGTTCCTCAAGAAACCAAGATGCAAT 459  
Qy 361 GGGAAATGCTCACTGTTCACTTACCTTAAAAAACAATGTAAAGATGCACTGATACCAT 420  
Db 460 GGGAAATGCTCACTGTTCACTTACCTTAAAAAACAATGTAAAGATGCACTGATACCAT 519  
Qy 421 AAAAGAAACCAAAAGAGGGTGTAGTGTCTGCTACGATATGAGTGGCGCTTCCAGTT 480  
Db 520 AAAAGAAACCAAAAGAGGGTGTAGTGTCTGCTACGATATGAGTGGCGCTTCCAGTT 579  
Qy 481 TACGAATGATCAAGCAACAGATGCTCTTAAATATGGAACGAAGATTAAGCGGTT 540  
Db 580 TACGAATGATCAAGCAACAGATGCTCTTAAATATGGAACGAAGATTAAGCGGTT 639  
Qy 541 GTTCAATATTTACTTAAAAATGTGTAGCAATGTATGATTTATCATGTGAAAAAGTA 600  
Db 640 GTTCAATATTTACTTAAAAATGTGTAGCAATGTATGATTTATCATGTGAAAAAGTA 699  
Qy 601 GGAATGCTGAAAAATGAGAGATTTAAATGGCGCAAGATTTTGTATTTCTTAAAGCGAAGGC 660  
Db 700 GGAATGCTGAAAAATGAGAGATTTAAATGGCGCAAGATTTTGTATTTCTTAAAGCGAAGGC 759  
Qy 661 TCACCAAGGCAAGTAAATATATCCAAAGAGTCAAGATGATTAATATCATGGAACAAG 720  
Db 760 TCACCAAGGCAAGTAAATATATCCAAAGAGTCAAGATGATTAATATCATGGAACAAG 819  
Qy 721 GATTAAGAAACAAGCAAAACGCTTTATTTACTGGGAAAAATTAATAATGGCGAAAAATGAT 780  
Db 820 GATTAAGAAACAAGCAAAACGCTTTATTTACTGGGAAAAATTAATAATGGCGAAAAATGAT 879  
Qy 781 TTCAACAAGACGAGAAATGGAACGGGAGAAATTAACGTTTAAAAATCTTGAAGTGTG 840  
Db 880 TTCAACAAGACGAGAAATGGAACGGGAGAAATTAACGTTTAAAAATCTTGAAGTGTG 939  
Qy 841 TATATTTTGAAGAAGTAAAGCTCAAAATTAATGCAAGTAAATTAATGAATCAACAAAA 900  
Db 940 TATATTTTGAAGAAGTAAAGCTCAAAATTAATGCAAGTAAATTAATGAATCAACAAAA 999  
Qy 901 ACACCAATTTACAATTTGAGCAAAACAATCAAAACCTGTGTAAGAAAAACAGTCAAAAATGAT 960  
Db 1000 ACACCAATTTACAATTTGAGCAAAACAATCAAAACCTGTGTAAGAAAAACAGTCAAAAATGAT 1059  
Qy 961 ACCTCTAAAGTTGATTAACCAACAACAAGCTTGAATGATTAAGATGTGCAATGGCGAA 1020  
Db 1060 ACCTCTAAAGTTGATTAACCAACAACAAGCTTGAATGATTAAGATGTGCAATGGCGAA 1119  
Qy 1021 AAAATTAATATCAATTTCTGTAATATTTCAATGGGAGTTGACAGCAAAAGAGCGGAC 1080  
Db 1120 AAAATTAATATCAATTTCTGTAATATTTCAATGGGAGTTGACAGCAAAAGAGCGGAC 1179  
Qy 1081 GCTAATTAATACGTCAATTTCAATTTAGTTATTAACATGATGACGCTTAACTTTTGAAT 1140

Db 1180 GCTAATTAATACGTCAATTTCAATTTAGTTATTAACATGATGACGCTTAACTTTGAT 1239  
Qy 1141 AACGTGACTTCTGGAGAGTATGCTTAATGCGTTATATATGATGGGAGATACGATATGCTCT 1200  
Db 1240 AACGTGACTTCTGGAGAGTATGCTTAATGCGTTATATGATGGGAGATACGATATGCTCT 1299  
Qy 1201 GAAATTTTCAAGTACGTAACAAGCAATGAGCTTCACTGTGCGCGTTAATCAACGATAT 1260  
Db 1300 GAAATTTTCAAGTACGTAACAAGCAATGAGCTTCACTGTGCGCGTTAATCAACGATAT 1359  
Qy 1261 ATTCCTACGCTTAACAACAGCGGCACTAAATTTGTTTACTTTATGATTTAAATGAA 1320  
Db 1360 ATTCCTACGCTTAACAACAGCGGCACTAAATTTGTTTACTTTATGATTTAAATGAA 1419  
Qy 1321 AAAGCAATTCCTACGAAGGCTTTTAAATATGAGGGAATGTTAATAGGTATACCGAC 1380  
Db 1420 AAAGCAATTCCTACGAAGGCTTTTAAATATGAGGGAATGTTAATAGGTATACCGAC 1479  
Qy 1381 GACCAAAACCAACCAACTGTTGAAGTTGTGACAGGTGGGAAAAGTTTCATTTAAAGTCAT 1440  
Db 1480 GACCAAAACCAACCAACTGTTGAAGTTGTGACAGGTGGGAAAAGTTTCATTTAAAGTCAT 1539  
Qy 1441 GCGGATGTGACAGCAACAAGCTTGGCGGAGCTTCTTTGTGTCTCGTGATCAAAAC 1500  
Db 1540 GCGGATGTGACAGCAACAAGCTTGGCGGAGCTTCTTTGTGTCTCGTGATCAAAAC 1599  
Qy 1501 AGCGACAGCAAAATTTATTTGAAATTCATGATGAAACAAGAAAGCAACCTTGGTGAAA 1560  
Db 1600 AGCGACAGCAAAATTTATTTGAAATTCATGATGAAACAAGAAAGCAACCTTGGTGAAA 1659  
Qy 1561 ACAAAAGCTGAAGCAACTACTTTTACAACAACGCGCTGATGATTAATGATATCAAGG 1620  
Db 1660 ACAAAAGCTGAAGCAACTACTTTTACAACAACGCGCTGATGATTAATGATATCAAGG 1719  
Qy 1621 CTTAAATACGTAATCTTATTTTAAAGAACTGTAGCTCCGTGATTAATGTCTTTGTA 1680  
Db 1720 CTTAAATACGTAATCTTATTTTAAAGAACTGTAGCTCCGTGATTAATGTCTTTGTA 1779  
Qy 1681 ACAAAATCGGATTAATTTGTGTGCAATGAAACAATGATGGAACAACGAAAACTAGTT 1740  
Db 1780 ACAAAATCGGATTAATTTGTGTGCAATGAAACAATGATGGAACAACGAAAACTAGTT 1839  
Qy 1741 TCACCAAGAAAGTACCAACAACAACAAGGTACTTAACCTTCAACAGTGGCAAGGA 1800  
Db 1840 TCACCAAGAAAGTACCAACAACAACAAGGTACTTAACCTTCAACAGTGGCAAGGA 1899  
Qy 1801 ATCTACGTTTACTTGAAGTGGCGCACTTGTCTACTTATTTGACGAGTCTACTTTGCT 1860  
Db 1900 ATCTACGTTTACTTGAAGTGGCGCACTTGTCTACTTATTTGACGAGTCTACTTTGCT 1959  
Qy 1861 AGACGTAGAAAAAGAAATGCT 1881  
Db 1960 AGACGTAGAAAAAGAAATGCT 1980

RESULT 5  
ABX61663  
ID ABX61663 strand: DNA; 1983 BP.  
XX  
AC ABX61663;  
XX  
DT 26-FEB-2003 (first entry)  
XX  
DE Enterococcus faecalis EF040 polymnucleotide #109.  
XX  
KW EF040; gene; ds; immunostimulant; antibacterial; gene mapping.  
XX  
OS Enterococcus faecalis.  
XX  
PN US6448043-B1.  
XX  
PD 10-SEP-2002.  
XX



PF 04-MAY-1998; 98US-00071035.  
XX  
PR 06-MAY-1997; 97US-0044031P.  
PR 16-MAY-1997; 97US-004655P.  
PR 14-NOV-1997; 97US-0066009P.  
PR 14-NOV-1997; 97US-0066099P.  
XX  
PA (HUMA-) HUMAN GENOME SCI INC.  
PI Choi GH, Bailey C, Hromockyj A, Kunsch CA,  
DR MPI; 2003-089120/08.  
XX P-PSDB; ABU13616.  
PT New EF040 polypeptides and polynucleotides from *Enterococcus faecalis*,  
PT useful for generating an immune response against *E. faecalis* and other  
XX *Enterococcus* species, and as vaccines against other bacterial genera.  
XX  
PS Example 1; Col 113-116; 146pp; English.  
XX  
CC The invention relates to polynucleotide fragments of a gene from  
CC *Enterococcus faecalis*, EF040, and the polypeptides encoded by them. The  
CC polypeptides are useful in detecting *E. faecalis*, as epitope tags, as  
CC molecular weight markers on SDS-PAGE gels or for molecular sieve gel  
CC filtration columns, in generating antibodies that specifically bind to  
CC the *E. faecalis* polypeptides, in generating an immune response against *E.*  
CC *faecalis* and other *Enterococcus* species and as vaccines against other  
CC bacterial genera. The polynucleotides are useful as probes for gene  
CC mapping and for identifying *E. faecalis* in biological samples. Sequences  
CC ABX6155-ABX61802 represent EF040 polynucleotides of the invention. Note:  
CC The sequence data for this patent can also be obtained from USPTO at  
XX [seqdata.uspto.gov/sequence.html](http://seqdata.uspto.gov/sequence.html)  
SQ Sequence 1983 BP; 729 A; 330 C; 426 G; 498 T; 0 U; 0 Other;

Query Match 100.0%; Score 1881; DB 8; Length 1983;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1881; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAGCAATTAAAAAGTTGGTACACCGTTAGTACCTTGTACTTAATTTGGCACTT 60  
DB 100 ATGAGCAATTAAAAAGTTGGTACACCGTTAGTACCTTGTACTTAATTTGGCACTT 159  
QY 61 TTCACAGGTATTTAGGCAACAACGTGATTTGCAAGAAATGGGAGACGCAAG 120  
DB 160 TTCACAGGTATTTAGGCAACAACGTGATTTGCAAGAAATGGGAGACGCAAG 219  
QY 121 CTCGTGATTCACAAAAAGAAATGACGATTTACAGATCCGCTTATTCAAAATAGCGG 180  
DB 220 CTCGTGATTCACAAAAAGAAATGACGATTTACAGATCCGCTTATTCAAAATAGCGG 279  
QY 181 AAAAGAAATGACGAGTTGATTAATATCAAGAGTGGCAGATGTGACGTTTACTTAT 240  
DB 280 AAAAGAAATGACGAGTTGATTAATATCAAGAGTGGCAGATGTGACGTTTACTTAT 339  
QY 241 AACGTGACGAAGAAATTTTACGAGCAACGAGCGGCGGCAAGCCTTGATGACGTA 300  
DB 340 AACGTGACGAAGAAATTTTACGAGCAACGAGCGGCGGCAAGCCTTGATGACGTA 399  
QY 301 CAAGCTGTCCAAAGTTTAACTCTGCGGAAACCTGTGCTCAAGGAACCAACGATGCAAT 360  
DB 400 CAAGCTGTCCAAAGTTTAACTCTGCGGAAACCTGTGCTCAAGGAACCAACGATGCAAT 459  
QY 361 GGGAAATGCTACTGTTCACTTAACTTAATAAACAATATGTAAAGATCAAGTATACCAT 420  
DB 460 GGGAAATGCTACTGTTCACTTAACTTAATAAACAATATGTAAAGATCAAGTATACCAT 519  
QY 421 AAAAGAAACCAAAAGAGGTGTAGTGTGCTGCTACGAATATGTGTGTGCGTCCAGTT 480  
DB 520 AAAAGAAACCAAAAGAGGTGTAGTGTGCTGCTACGAATATGTGTGTGCGTCCAGTT 579  
QY 481 TAGAAATGATCAAGCAACAGATGCTTCTATTAATATGAAACAGAAATTTAGCGTT 540

DB 580 TAGAAATGATCAAGCAACAGATGCTTCTATTAATATGAAACAGAAATTTAGCGTT 639  
QY 541 GTTCATATTTTATCTTAATAAATGTGTAGCCATGATGTAAGTTTACTGTGAAAAAGTA 600  
DB 640 GTTCATATTTTATCTTAATAAATGTGTAGCCATGATGTAAGTTTACTGTGAAAAAGTA 699  
QY 601 GGAACCTGTGAAAAATGAAGATTTAATGGCGCAATTTGTTATTTCTAAAGCAAGGC 660  
DB 700 GGAACCTGTGAAAAATGAAGATTTAATGGCGCAATTTGTTATTTCTAAAGCAAGGC 759  
QY 661 TCACGAGCAAGTAAATATATATCAAGAGTCAAGATGATATATATATATATATATATAT 720  
DB 760 TCACGAGCAAGTAAATATATATATCAAGAGTCAAGATGATATATATATATATATATAT 819  
QY 721 GATTAAGAAACAGCAAAAGCCTTATTTACTGGGAAAAGTTATGAAATTTGGCGAAATGAT 780  
DB 820 GATTAAGAAACAGCAAAAGCCTTATTTACTGGGAAAAGTTATGAAATTTGGCGAAATGAT 879  
QY 781 TTCAGAGAGCAGAGATGGAACGAGGAGATTAAAGTTTAAATCTTGAGGTTGTTG 840  
DB 880 TTCAGAGAGCAGAGATGGAACGAGGAGATTAAAGTTTAAATCTTGAGGTTGTTG 939  
QY 841 TATATTTTGAAGAGTAAAGCTCCAAATATATGAGAAATTAATGAAATCAACAGAA 900  
DB 940 TATATTTTGAAGAGTAAAGCTCCAAATATATGAGAAATTAATGAAATCAACAGAA 999  
QY 901 ACACCTTTTACATTTGAGCAAAATCAACACCTGTTGAAAAACAGTCAAAATGAT 960  
DB 1000 ACACCTTTTACATTTGAGCAAAATCAACACCTGTTGAAAAACAGTCAAAATGAT 1059  
QY 961 ACCTTAAAGTTGATTAACCAACAGAGCTTATGATGATGATGATGATGATGATGATGAT 1020  
DB 1060 ACCTTAAAGTTGATTAACCAACAGAGCTTATGATGATGATGATGATGATGATGATGAT 1119  
QY 1021 AAAATTTAAATCAAAATTTCTGTAATATTCATTTGGGATTCAGACAAAGAGCGAC 1080  
DB 1120 AAAATTTAAATCAAAATTTCTGTAATATTCATTTGGGATTCAGACAAAGAGCGAC 1179  
QY 1081 GCTAATTAATAGTCAAAATTTCAATTTAGTTGATTAACATGATGACCTTAATCTTTGAT 1140  
DB 1180 GCTAATTAATAGTCAAAATTTCAATTTAGTTGATTAACATGATGACCTTAATCTTTGAT 1239  
QY 1141 AACGTGATTCGGAAGATGATGCTTATGCTTATATGATGATGATGATGATGATGATGAT 1200  
DB 1240 AACGTGATTCGGAAGATGATGCTTATGCTTATATGATGATGATGATGATGATGATGAT 1299  
QY 1201 GAAATTTATCAAGTGACTGAACAGCAATGCGTCACTGTGCGCTTAATCCAGCGTAT 1260  
DB 1300 GAAATTTATCAAGTGACTGAACAGCAATGCGTCACTGTGCGCTTAATCCAGCGTAT 1359  
QY 1261 ATTCTGACGCTAACACGAGCGGCACTTAATTTGTTTACTTTATGCAATTTAAATGAA 1320  
DB 1360 ATTCTGACGCTAACACGAGCGGCACTTAATTTGTTTACTTTATGCAATTTAAATGAA 1419  
QY 1321 AAAAGAGATCTTACGAAAGGCTTTAAATATGAGCGAATGTTGATTAACGCTTACGAC 1380  
DB 1420 AAAAGAGATCTTACGAAAGGCTTTAAATATGAGCGAATGTTGATTAACGCTTACGAC 1479  
QY 1381 GACCAAAACCAACCAACTGTGTAAGTTGACAGGTGGGAAACGTTTCAATTAAGTCAT 1440  
DB 1480 GACCAAAACCAACCAACTGTGTAAGTTGACAGGTGGGAAACGTTTCAATTAAGTCAT 1539  
QY 1441 GGGGATGTGACAGCGACCAAGCCTTGGCGGAGCTTCTTGTGCTCGGTATCAAAAC 1500  
DB 1540 GGGGATGTGACAGCGACCAAGCCTTGGCGGAGCTTCTTGTGCTCGGTATCAAAAC 1599  
QY 1501 AGGCAACAGCAAAATTAATTTGAAAAATCGATGAAACAAAGAAAGCAACTTGGGTGAAA 1560  
DB 1600 AGGCAACAGCAAAATTAATTTGAAAAATCGATGAAACAAAGAAAGCAACTTGGGTGAAA 1659  
QY 1561 ACAAAAGCTGAAGCAACTTATTTTCAACAACGCGTATGATGATGATGATGATGATGATGAT 1620  
DB 1660 ACAAAAGCTGAAGCAACTTATTTTCAACAACGCGTATGATGATGATGATGATGATGATGAT 1719



```

QY 1021 AAAATTAATCAATTTCTGTAATATTCATTTGGGGATGGACAAAGAGCGAC 1080
DB 1120 AAAATTAATCAATTTCTGTAATATTCATTTGGGGATGGACAAAGAGCGAC 1179
QY 1081 GCTAATTAATCAATTTCTGTAATATTCATTTGGGGATGGACAAAGAGCGAC 1140
DB 1180 GCTAATTAATCAATTTCTGTAATATTCATTTGGGGATGGACAAAGAGCGAC 1239
QY 1141 AACGGAATCTTGGAGAGATATGCTTATATGATGGGGATACAGTATGCTCT 1200
DB 1240 AACGGAATCTTGGAGAGATATGCTTATATGATGGGGATACAGTATGCTCT 1299
QY 1201 GAAATTAATCAATTTCTGTAATATTCATTTGGGGATGGACAAAGAGCGAC 1260
DB 1300 GAAATTAATCAATTTCTGTAATATTCATTTGGGGATGGACAAAGAGCGAC 1359
QY 1261 ATTCTAGCTTAACACGAGCGGACACTTAAATTCGTTTACTTTATGCAATTAATGA 1320
DB 1360 ATTCTAGCTTAACACGAGCGGACACTTAAATTCGTTTACTTTATGCAATTAATGA 1419
QY 1321 AAAGCAATCTTACGAAAGGCTTTAAATATGAGCGAATGTTATACGCTCATCCGAC 1380
DB 1420 AAAGCAATCTTACGAAAGGCTTTAAATATGAGCGAATGTTATACGCTCATCCGAC 1479
QY 1381 GACCAACACGACCACTGTTGAGTGTGACAGGTGGGAAAGCTTCAATTAAGTCGAT 1440
DB 1480 GACCAACACGACCACTGTTGAGTGTGACAGGTGGGAAAGCTTCAATTAAGTCGAT 1539
QY 1441 GCGGATGTGACAGCAACAGCCTTGGCGGAGCTTCCCTTTCGTCCTGATCAAAAC 1500
DB 1540 GCGGATGTGACAGCAACAGCCTTGGCGGAGCTTCCCTTTCGTCCTGATCAAAAC 1599
QY 1501 AGCGACACGCAATTTATTTGAAATCGATGAAACAGAAAGCAAGCACTTGGTGAA 1560
DB 1600 AGCGACACGCAATTTATTTGAAATCGATGAAACAGAAAGCAAGCACTTGGTGAA 1659
QY 1561 ACAAAAGCTGAAGCACTTCTTCAACAGCGGCTGATGATTTGTTATATCAAGGG 1620
DB 1660 ACAAAAGCTGAAGCACTTCTTCAACAGCGGCTGATGATTTGTTATATCAAGGG 1719
QY 1621 CTTAAATACGCTTATTTATTAAGAAACCTGAGCTCTGATGATTTATCTTGTGA 1680
DB 1720 CTTAAATACGCTTATTTATTAAGAAACCTGAGCTCTGATGATTTATCTTGTGA 1779
QY 1681 ACAATTCGATTTGATTTGTGCTCAATGAAACATATGAGCAACAGAAACCTAGTT 1740
DB 1780 ACAATTCGATTTGATTTGTGCTCAATGAAACATATGAGCAACAGAAACCTAGTT 1839
QY 1741 TCACAGAAAGATGCAACAAAGCAAGAGTACCTTCAACAGGTGGCAAGGA 1800
DB 1840 TCACAGAAAGATGCAACAAAGCAAGAGTACCTTCAACAGGTGGCAAGGA 1899
QY 1801 ATCTAGTTTACTTGAAGTGGCAGCTTGTACTTATTTGACAGAGTCTACTTGTCT 1860
DB 1900 ATCTAGTTTACTTGAAGTGGCAGCTTGTACTTATTTGACAGAGTCTACTTGTCT 1959
QY 1861 AGACGTAGAAAGAAATGCT 1881
DB 1960 AGACGTAGAAAGAAATGCT 1980

```

## RESULT 7

AAK12982 ID AAK12982 standard; DNA; 15614 BP.

AAK12982;

19-MAR-1999 (first entry)

Enterococcus faecalis genome contig SEQ ID NO:45.

Enterococcus faecalis; contig; detection; Enterococcal infection;

vaccine; attenuation; computer readable medium; ds.

```

XX XX Enterococcus faecalis.
OS WO980555-A2.
PN 12-NOV-1998.
PD 04-MAY-1998; 98WO-US008985.
PF 06-MAY-1997; 97US-0044031P.
PR 16-MAY-1997; 97US-0046555P.
PR 14-NOV-1997; 97US-0065009P.
XX (HUMA-) HUMAN GENOME SCI INC.
XX Kunesh CA, Dillon PV, Barash SC,
PI WPI; 1999-045171/04.
XX DR
XX PT New isolated Enterococcus faecalis polynucleotides and polypeptides -
XX used to develop products for the detection of Enterococcus and for use in
XX vaccines for prevention or attenuation of Enterococcus infection.
XX PS Claim 1; Page 419-427; 2084dp; English.
XX CC A computer readable medium has been developed which has recorded on it
CC 982 nucleotide sequences isolated from the Enterococcus faecalis genome.
CC AAK12982 to AAK13919 represent these nucleotide sequences which are
CC primary nucleotide sequences, also known as contigs. The computer-based
CC system can identify fragments of the Enterococcus faecalis genome with
CC commercial importance. The products can be used to detect the presence of
CC Enterococcus faecalis in samples. They can also be used for diagnosing
CC Enterococcal infection in an animal and monitoring progression of
CC disease, and for identifying agents which can be used to modulate the
CC growth or pathogenicity of Enterococcus faecalis, or another related
CC organism, in vivo or in vitro. In particular the polypeptides encoded by
CC the Enterococcus faecalis nucleotide sequences can be used in vaccines to
CC prevent or attenuate an Enterococcal infection
XX SQ Sequence 15614 BP; 5535 A; 2567 C; 3169 G; 4326 T; 0 U; 17 Other;
Query Match 100.0%; Score 1881; DB 2; Length 15614;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1881; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGAAGCAATTAATAAAGTTGTGACACCGTTAGCTTGTACTTAATTTGCCACTT 60
DB 12164 ATGAAGCAATTAATAAAGTTGTGACACCGTTAGCTTGTACTTAATTTGCCACTT 12223
QY 61 TTCACAGTGTATTTGGGACACAACTGCAATTTGCAAGAAATGGGGAGCGCACAG 120
DB 12224 TTCACAGTGTATTTGGGACACAACTGCAATTTGCAAGAAATGGGGAGCGCACAG 12283
QY 121 CTCGTGATTTCAAAAAGAAATGACGAGTTTACAGATCCGCTTATTTCAAATAGCGGG 180
DB 12284 CTCGTGATTTCAAAAAGAAATGACGAGTTTACAGATCCGCTTATTTCAAATAGCGGG 12343
QY 181 AAAAGAAATGAGCGAGTTGATTAATATCAAGGACTGGCAGATGTGACGTTAGTATAT 240
DB 12344 AAAAGAAATGAGCGAGTTGATTAATATCAAGGACTGGCAGATGTGACGTTAGTATAT 12403
QY 241 AACGTGACGAACGAATTTTACAGACAGAGCGGACGACGAGCTTGATGACAGTAA 300
DB 12404 AACGTGACGAACGAATTTTACAGACAGAGCGGACGACGAGCTTGATGACAGTAA 12463
QY 301 CAAGCTGTCCAAAGTTTAACTCTGGGAAACCTGTGCTCAAGGAACCAAGCATCAAT 360
DB 12464 CAAGCTGTCCAAAGTTTAACTCTGGGAAACCTGTGCTCAAGGAACCAAGCATCAAT 12523
QY 361 GGGAAATGCTACTGTTCACTTAATAAACAACAAATGTTAAAGATGACAGTATACAT 420
DB 12524 GGGAAATGCTACTGTTCACTTAATAAACAACAAATGTTAAAGATGACAGTATACAT 12583

```

421 AAAGAAAGCAAAAGAGGTGTAAGTGTGCTAGCAATATGATGATGGCGTTCCAGTT 480  
12584 AAAGAAAGCAAAAGAGGTGTAAGTGTGCTAGCAATATGATGATGGCGTTCCAGTT 12643  
481 TAAGCAATGATCAAGCAAAAGAGGTGTAAGTGTGCTAGCAATATGATGATGGCGTT 540  
12644 TAAGCAATGATCAAGCAAAAGAGGTGTAAGTGTGCTAGCAATATGATGATGGCGTT 12703  
541 GTTCAATATTTATCCCTAAATATGATGATGGCGTTAGCAATATGATGATGGCGTT 600  
12704 GTTCAATATTTATCCCTAAATATGATGATGGCGTTAGCAATATGATGATGGCGTT 12763  
601 GGAATGCTGTAATATGATGATGGCGTTAGCAATATGATGATGGCGTTAGCAATATG 660  
12764 GGAATGCTGTAATATGATGATGGCGTTAGCAATATGATGATGGCGTTAGCAATATG 12823  
661 TCACCAAGCAAGTAATATATCCAAAGAGTCAAGATGATGATGATGATGATGATGATG 720  
12824 TCACCAAGCAAGTAATATATCCAAAGAGTCAAGATGATGATGATGATGATGATGATG 12883  
721 GATTAAGCAAGCAAAAGAGGTGTAAGTGTGCTAGCAATATGATGATGGCGTTAGCA 780  
12884 GATTAAGCAAGCAAAAGAGGTGTAAGTGTGCTAGCAATATGATGATGGCGTTAGCA 12943  
781 TTCACAGAAAGCAAGATGATGATGGCGTTAGCAATATGATGATGGCGTTAGCAATG 840  
12944 TTCACAGAAAGCAAGATGATGATGGCGTTAGCAATATGATGATGGCGTTAGCAATG 13003  
841 TATATTTTGAAGAAAGTAAGCTCCAAATATGCAATATGCAATATGCAATATGCAAT 900  
13004 TATATTTTGAAGAAAGTAAGCTCCAAATATGCAATATGCAATATGCAATATGCAAT 13063  
901 ACAACATTTCAATTTGAAGCAAAAGCAATGCAATGCTGTAATGCAATGCTGTAATG 960  
13064 ACAACATTTCAATTTGAAGCAAAAGCAATGCAATGCTGTAATGCAATGCTGTAATG 13123  
961 ACCTCTAAAGTGTATTAACCAACCAAGCTTATGATGATGATGATGATGATGATGATG 1020  
13124 ACCTCTAAAGTGTATTAACCAACCAAGCTTATGATGATGATGATGATGATGATGATG 13183  
1021 AAAATTTAATATCAATTTCTGTAATATTCATTTGGGATGTCAGACAAAGAGCGAC 1080  
13184 AAAATTTAATATCAATTTCTGTAATATTCATTTGGGATGTCAGACAAAGAGCGAC 13243  
1081 GCTAATTAATAGTCAATTTCAATTTAGTGAATTAACAGATGATGATGATGATGATG 1140  
13244 GCTAATTAATAGTCAATTTCAATTTAGTGAATTAACAGATGATGATGATGATGATG 13303  
1141 AACGTACTTCTGAGAGATGCTTATGCTTATATGATGATGATGATGATGATGATGATG 1200  
13304 AACGTACTTCTGAGAGATGCTTATGCTTATATGATGATGATGATGATGATGATGATG 13363  
1201 GAAAATTTATCAAGTGAATGCAAGCAATGCTTATGATGATGATGATGATGATGATG 1260  
13364 GAAAATTTATCAAGTGAATGCAAGCAATGCTTATGATGATGATGATGATGATGATG 13423  
1261 ATTCTAGCTTAAACCAAGCGGCAACATTAATTTGCTTATGATGATGATGATGATG 1320  
13424 ATTCTAGCTTAAACCAAGCGGCAACATTAATTTGCTTATGATGATGATGATGATG 13483  
1321 AAAGCAATCTTCAAGAAAGCTTTTAAATGAGCGCAATGCTTAAAGCTTAAAGCTT 1380  
13484 AAAGCAATCTTCAAGAAAGCTTTTAAATGAGCGCAATGCTTAAAGCTTAAAGCTT 13543  
1381 GACCAAAACCAACCACTGTTGAGTGTGACAGTGTGGAACGTTTCAATTAAGTGCAT 1440  
13544 GACCAAAACCAACCACTGTTGAGTGTGACAGTGTGGAACGTTTCAATTAAGTGCAT 13603  
1441 GCGGATGATGACAGCGCAACAGCTTGGGCGGAGCTTCTGCTGCTGCTGATGCAAAAC 1500  
13604 GCGGATGATGACAGCGCAACAGCTTGGGCGGAGCTTCTGCTGCTGCTGATGCAAAAC 13663  
1501 AGGACACAGCAAAATTTTGAATGATGAAACCAAGCAAGCAAACTTGGGTGAAA 1560

13664 AGGACACAGCAAAATTTTGAATGATGAAACCAAGCAAGCAAACTTGGGTGAAA 13723  
1561 ACMAAGCTGAAAGCACTACTTTTACCAACCAAGCTGATGATGATGATGATGATGATG 1620  
13724 ACMAAGCTGAAAGCACTACTTTTACCAACCAAGCTGATGATGATGATGATGATGATG 13783  
1621 CTTAATATGCGTACTTATTTATGAAAGAACTGATGCTGCTGATGATGATGATGATG 1680  
13784 CTTAATATGCGTACTTATTTATGAAAGAACTGATGCTGCTGATGATGATGATGATG 13843  
1681 ACMAATGCGATTTGAATTTGCTGATGAAACATCATATGCAACCAAGAAACCTAGTT 1740  
13844 ACMAATGCGATTTGAATTTGCTGATGAAACATCATATGCAACCAAGAAACCTAGTT 13903  
1741 TCACCAAGAAAGTACCAACCAAGTACTTATCTTCAACAGTGGCAAGGA 1800  
13904 TCACCAAGAAAGTACCAACCAAGTACTTATCTTCAACAGTGGCAAGGA 13963  
1801 ATCTAGTTTACTTGAAGTGGCGCAGCTTGTACTTATTTGACAGTCTACTTTGCT 1860  
13964 ATCTAGTTTACTTGAAGTGGCGCAGCTTGTACTTATTTGACAGTCTACTTTGCT 14023  
1861 AGACGTAGAAAGAAATGCT 1881  
14024 AGACGTAGAAAGAAATGCT 14044

RESULT 8  
ABS98777  
ID ABS98777 standard; DNA; 15614 BP.  
XX  
XX  
AC ABS98777;  
XX  
DT 18-DEC-2002 (first entry)  
XX  
DE Enterococcus faecalis contig sequence #45.  
XX  
KW Computer readable medium; Enterococcus faecalis; microbe; growth;  
XX pathogenicity; vaccine; resistance; Enterococcal infection; commercial;  
KW therapeutic; industrial; fermenting; sugar source; metabolite; vaccine;  
KW biotech; technology; antibacterial; modulator of nucleic acid expression;  
KW config; ds.  
XX  
OS Enterococcus faecalis.  
XX  
PN US2002120116-A1.  
XX  
PD 29-AUG-2002.  
XX  
PF 04-MAY-1998; 98US-00070927.  
XX  
PR 04-MAY-1998; 98US-00070927.  
XX  
PA (KUNS/) KUNSCH C A.  
XX (DILL/) DILLON P J.  
XX (BARA/) BARASH S.  
XX  
PI Kunsch CA, Dillon PJ, Barash S;  
XX  
DR WPI; 2002-750065/81.  
XX  
PT Computer readable medium having recorded on it a Enterococcus faecalis  
XX nucleotide sequence useful for detecting diseases related to Enterococcus  
XX infections in animals.  
XX  
XX Claim 1; Page; 119pp; English.  
XX  
CC The present invention relates to a new computer readable medium with an  
XX Enterococcus faecalis nucleotide sequence. The invention is useful to  
XX diagnose the presence of E.faecalis in a sample or determining the  
XX presence of a specific microbe in a sample. The invention is also useful  
XX for modulating the growth or pathogenicity of E.faecalis, in a vaccine to

CC confer resistance to Enterococcal infection, for commercial, therapeutic  
CC and industrial purposes, and for fermenting a particular sugar source or  
CC to produce a particular metabolite. The invention is useful for detecting  
CC diseases related to Enterococcus infections in animals, and for detecting  
CC E. faecalis using bioclip technology. The present nucleic acid sequence  
CC represents an Enterococcus faecalis contig DNA sequence of the invention.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification but was obtained in electronic format directly from USPTO  
CC at http://seqdata.uspto.gov

XX Sequence 15614 BP; 5535 A; 2567 C; 3169 G; 4326 T; 0 U; 17 Other;

Query Match 100.0%; Score 1881; DB 6; Length 15614;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1881; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 ATGAGCAATTTAAAAAGTTGGTACACCGTTAGTACCTGTGTACTAAATTTTGCACCT 60
DB 12164 ATGAGCAATTTAAAAAGTTGGTACACCGTTAGTACCTGTGTACTAAATTTTGCACCT 12223
QY 61 TTCACAGGTGTATTAGGACACAACTGCAATTTGCAAGAAAAATGGGAGAGCCGACAG 120
DB 12224 TTCACAGGTGTATTAGGACACAACTGCAATTTGCAAGAAAAATGGGAGAGCCGACAG 12283
QY 121 CTCGTGATTCACAAAAAGAAAATGACGGAATTTACAGATCCGCTTATTCAAAATAGCGGG 180
DB 12284 CTCGTGATTCACAAAAAGAAAATGACGGAATTTACAGATCCGCTTATTCAAAATAGCGGG 12343
QY 181 AAAAGAAATGACGAGTTGATTAATATYCAAGACTGCGAAGATGTGACGTTTATGATTTAT 240
DB 12344 AAAAGAAATGACGAGTTGATTAATATYCAAGACTGCGAAGATGTGACGTTTATGATTTAT 12403
QY 241 AACGTGACGAAGAAATTTTACAGACGACGCGGCGGCGCAAGCTTGATGACGCTAA 300
DB 12404 AACGTGACGAAGAAATTTTACAGACGACGCGGCGGCGCAAGCTTGATGACGCTAA 12463
QY 301 CAAGCTGTCCAAAGTTTAACTCTGGGGAACCTGTGCTCAAGGAACACCGATGCAAT 360
DB 12464 CAAGCTGTCCAAAGTTTAACTCTGGGGAACCTGTGCTCAAGGAACACCGATGCAAT 12523
QY 361 GGGAAATGCTACTGTTCAAGTTACCTTAAAAAACAATGATGTAAGTCACTGATACAT 420
DB 12524 GGGAAATGCTACTGTTCAAGTTACCTTAAAAAACAATGATGTAAGTCACTGATACAT 12583
QY 421 AAAAGAAACCAAAAGAGGTGTAGTGTGCTGACGAATATGTGTGCGCTTCCAGTT 480
DB 12584 AAAAGAAACCAAAAGAGGTGTAGTGTGCTGACGAATATGTGTGCGCTTCCAGTT 12643
QY 481 TACGAATATGATCAAGCAACAGATGCTCTTAAATATGGAACAGAAATTAAGCGGT 540
DB 12644 TACGAATATGATCAAGCAACAGATGCTCTTAAATATGGAACAGAAATTAAGCGGT 12703
QY 541 GTTCATATTTATCTTAAAAATGTGTAGCAATGATGTAAGTTTACATGTGAAAAAGTA 600
DB 12704 GTTCATATTTATCTTAAAAATGTGTAGCAATGATGTAAGTTTACATGTGAAAAAGTA 12763
QY 601 GGAATGCTGAAAAAAGAGATTTAAATGGCGCAAGATTTGTTTCTTAAAGCGAAGCC 660
DB 12764 GGAATGCTGAAAAAAGAGATTTAAATGGCGCAAGATTTGTTTCTTAAAGCGAAGCC 12823
QY 661 TCACCAAGGACAGTAAAAATATATCCAAAGAGTCAAAAGATGATTTATATCATGACACAG 720
DB 12824 TCACCAAGGACAGTAAAAATATATCCAAAGAGTCAAAAGATGATTTATATCATGACACAG 12883
QY 721 GATTAAGAACAGCAAAACGCTTTTATATCTGAGAAAAAGTTATGAATTTGGCAAAATGAT 780
DB 12884 GATTAAGAACAGCAAAACGCTTTTATATCTGAGAAAAAGTTATGAATTTGGCAAAATGAT 12943
QY 781 TTCACAGAGCAGAGATGGAACGGGAGATTTAAACGTTAAATCTTGAGGTGCTG 840
DB 12944 TTCACAGAGCAGAGATGGAACGGGAGATTTAAACGTTAAATCTTGAGGTGCTG 13003
QY 841 TATATTTTGAAGAAATGAAGCTCCAAATATATGCAATTAATGAATCAAAACAAA 900
```

```
DB 13004 TATATTTTGAAGAAATGAAGCTCCAAATATATGCAATTAATGAATCAAAACAAA 13063
QY 901 ACACATTTTACATTTAGAACCAACATCAACACCTGTGAAAAAACAAGCAAAATGAT 960
DB 13064 ACACATTTTACATTTAGAACCAACATCAACACCTGTGAAAAAACAAGCAAAATGAT 13123
QY 961 ACCTCTAAAGTTGATTAACCAACACAGCTTATGATGTTAAAGATGTGGCAATTTGGCGAA 1020
DB 13124 ACCTCTAAAGTTGATTAACCAACACAGCTTATGATGTTAAAGATGTGGCAATTTGGCGAA 13183
QY 1021 AAAATTAATATCAAAATTTCTGTAATATATCCATTTGGGATTTGACAGAAAGAGCGAC 1080
DB 13184 AAAATTAATATCAAAATTTCTGTAATATATCCATTTGGGATTTGACAGAAAGAGCGAC 13243
QY 1081 GCTAATTAATACGTCAAAATTCATTTAGTTGATTAACATGATGACGCTTAACTTTGAT 1140
DB 13244 GCTAATTAATACGTCAAAATTCATTTAGTTGATTAACATGATGACGCTTAACTTTGAT 13303
QY 1141 AACGTGACTTCTGAGAGATGATGCTTATATGATATGAGGGGATTCAGTATGCTCCT 1200
DB 13304 AACGTGACTTCTGAGAGATGATGCTTATATGATATGAGGGGATTCAGTATGCTCCT 13363
QY 1201 GAAATTTATCAAGTACTGAACAGCAATGCGCTTCACTGCGCGCTTATCCAGCGTAT 1260
DB 13364 GAAATTTATCAAGTACTGAACAGCAATGCGCTTCACTGCGCGCTTATCCAGCGTAT 13423
QY 1261 ATTCTAGCGCTAACACAGCGCGGACACTTAAATTCGTTTACTTTATGCAATTTAAATGA 1320
DB 13424 ATTCTAGCGCTAACACAGCGCGGACACTTAAATTCGTTTACTTTATGCAATTTAAATGA 13483
QY 1321 AAAGCAGATCTTACGAAAGGCTTTTAAATATGAGGCGAAATGTTGATTAACGCTATACCAC 1380
DB 13484 AAAGCAGATCTTACGAAAGGCTTTTAAATATGAGGCGAAATGTTGATTAACGCTATACCAC 13543
QY 1381 GACCAAAACACACCAACTGTTGAAGTTGACAGGTGGGAAACGTTTCAATTAAGTCGAT 1440
DB 13544 GACCAAAACACACCAACTGTTGAAGTTGACAGGTGGGAAACGTTTCAATTAAGTCGAT 13603
QY 1441 GCGGATGTGACAGGACACACAGCCTTTGGCGGAGCTTCTTTTGTGCTCGGTATCAAAAC 1500
DB 13604 GCGGATGTGACAGGACACACAGCCTTTGGCGGAGCTTCTTTTGTGCTCGGTATCAAAAC 13663
QY 1501 AGCGACACAGCAATTTTGAATTCGATGAAACACAGAAAGCGACCACTTGGGTGAA 1560
DB 13664 AGCGACACAGCAATTTTGAATTCGATGAAACACAGAAAGCGACCACTTGGGTGAA 13723
QY 1561 ACAAAAGCTGAAGCAACTTCTTTTCAACAAAGGCTGATGATTTGTTGATATCAAGGG 1620
DB 13724 ACAAAAGCTGAAGCAACTTCTTTTCAACAAAGGCTGATGATTTGTTGATATCAAGGG 13783
QY 1621 CTTAAATACGGTACTTATTTTGAAGAAACGTGAGCTCTGATGATTTATGCTTTGTTA 1680
DB 13784 CTTAAATACGGTACTTATTTTGAAGAAACGTGAGCTCTGATGATTTATGCTTTGTTA 13843
QY 1681 ACAATTCGATTTGATTTTGTGCTCAATGAAACATATATGCAACCAAGAAACCTTACT 1740
DB 13844 ACAATTCGATTTGATTTTGTGCTCAATGAAACATATATGCAACCAAGAAACCTTACT 13903
QY 1741 TCACCAAGAAAGTACCAACCAACCAAGGTAACCTTACCTTCAACAGGTGGCAAGGA 1800
DB 13904 TCACCAAGAAAGTACCAACCAACCAAGGTAACCTTACCTTCAACAGGTGGCAAGGA 13963
QY 1801 ATCTACGTTTACTTGAAGATGGCGAGCTTGTCTACTTATTTGACAGAGCTTACTTTGCT 1860
DB 13964 ATCTACGTTTACTTGAAGATGGCGAGCTTGTCTACTTATTTGACAGAGCTTACTTTGCT 14023
QY 1861 AGAGCTAGAAAAAGAAATGCT 1881
DB 14024 AGAGCTAGAAAAAGAAATGCT 14044
```

RESULT 9

AD084888  
ID AD084888 standard; DNA; 1875 BP.  
XX  
AC AD084888;  
XX  
DT 29-JUL-2004 (first entry)  
XX  
DE E faecalis surface anchored LPXTG protein gene SeqID12.  
XX  
XX LPXTG; cell wall-anchored surface protein; Gram positive bacterium;  
XX extracellular matrix molecule; sequence database; C-terminal;  
XX immunoglobulin-like fold region; Ig-like fold region; antibacterial;  
XX vaccine; gene therapy; infection; medical device; prosthesis;  
XX premature newborn; AIDS; debilitated cancer; bone marrow transplantation;  
XX gene; ds.  
OS Enterococcus faecalis.  
XX  
XX  
XX Key Location/Qualifiers  
XX CDS 1..1875  
XX /tag= a  
XX /product= "E faecalis surface anchored LPXTG protein"  
XX /partial  
XX /note= "No stop codon"  
XX  
XX MO2004025416-A2.  
XX  
XX PD 25-MAR-2004.  
XX  
XX PF 15-SEP-2003; 2003WO-US028789.  
XX  
XX PR 13-SEP-2002; 2002US-0410303P.  
XX  
XX PA (TEXA ) UNIV TEXAS A & M SYSTEM.  
XX (INH1-) INHIBITEX INC.  
XX PA (UABR-) UAB RES FOUNDED.  
XX  
XX PI Hook M, Xu Y, Sillanpaa JV, Sthanam N, Ponnuraj K, Patti JM;  
PI Hutchins JT, Hall A;  
XX  
XX WPI: 2004-315684/29.  
DR P-PSDB; AD084889.  
XX  
XX  
XX Identifying LPXTG-containing cell wall-anchored surface proteins from  
XX Gram positive bacteria, for treating infection caused by the bacteria,  
XX comprises searching sequence information database for the sequence having  
XX LPXTG-motif.  
XX  
XX Claim 18; SEQ ID NO 12; 96bp; English.  
XX  
XX This invention relates to a novel method of identifying LPXTG-containing  
XX cell wall-anchored surface proteins from Gram positive bacteria that bind  
XX to an extracellular matrix molecule which comprises searching a database  
XX of sequence information for a putative protein sequence having the LPXTG-  
XX motif in its C-terminal region and analysing the sequence for the  
XX presence of one or more immunoglobulin (Ig)-like fold regions. The  
XX invention may be useful for the production of compounds with an  
XX antibacterial activity or for production of a vaccine. In addition the  
XX disclosed sequences may be useful for gene therapy. The antibody is  
XX useful for treating or preventing an infection of Gram-positive bacteria  
XX in a human or animal patient. The method and the proteins are useful in  
XX generating antibodies for treating and preventing the spread of  
XX infections of Gram positive bacteria, for interfering with, or inhibiting  
XX binding interactions by Gram positive bacteria, for monitoring the level  
XX of gram positive bacterial antigens, or antibodies recognizing the  
XX antigens in a human or animal patients suspected of containing the  
XX antigens or antibodies, in preventing or reducing infection of medical  
XX devices and prostheses caused by such organisms, and in treating or  
XX preventing infections in highly susceptible groups such as premature  
XX newborns, AIDS and debilitated cancer patients, and bone marrow  
XX transplantation. The present sequence is that of a gene which encodes a  
XX surface anchored LPXTG protein identified using the method of the  
XX invention.

XX  
SQ Sequence 1875 BP; 691 A; 317 C; 400 G; 467 T; 0 U; 0 Other;  
Query Match 98.8%; Score 1859; DB 12; Length 1875;  
Best Local Similarity 99.7%; Pred. No. 0;  
Matches 1875; Conservative 0; Mismatches 0; Indels 6; Gaps 1;  
QY 1 ATGAGCAATTAAAAAGTTTGTTGATACCCGTAGTACCTTGTCTAATTTTGGCACTT 60  
DB 1 ATGAGCAATTAAAAAGTTTGTTGATACCCGTAGTACCTTGTCTAATTTTGGCACTT 60  
QY 61 TTCAAGATGATTTATGAGCAACAACCTGATTTGCAAGAAATGGGAGAGCGACAG 120  
DB 61 TTCAAGATGATTTATGAGCAACAACCTGATTTGCAAGAAATGGGAGAGCGACAG 120  
QY 121 CTGCGATTTACAAAAGAAAATGACCGATTTTACAGATCCGCTTATTTCAAAATGCGGG 180  
DB 121 CTGCGATTTACAAAAGAAAATGACCGATTTTACAGATCCGCTTATTTCAAAATGCGGG 180  
QY 181 AAAAGAAATGAGCGAGTTTGTATTAATATGAGGACTGGCGATGTGACGTTTGTATTTAT 240  
DB 181 AAAAGAAATGAGCGAGTTTGTATTAATATGAGGACTGGCGATGTGACGTTTGTATTTAT 240  
QY 241 AACGTGAGAAAGCAATTTTACGAGCAACGAGCGGCAAGCGCTTGATGACACTAA 300  
DB 241 AACGTGAGAAAGCAATTTTACGAGCAACGAGCGGCAAGCGCTTGATGACACTAA 300  
QY 301 CAAGCTGTCCAAAGTTTAACTCTCGGAAACCTGTGTCTCAAGAAACCAACCGATGCCAAT 360  
DB 301 CAAGCTGTCCAAAGTTTAACTCTCGGAAACCTGTGTCTCAAGAAACCAACCGATGCCAAT 360  
QY 361 GGGAAATGTCATGTTCACTTACCTTAAATAAACAATATGTTAAAGATGATGATGAT 420  
DB 361 GGGAAATGTCATGTTCACTTACCTTAAATAAACAATATGTTAAAGATGATGATGAT 420  
QY 421 AAAAGAAACCAAAAGAGGTGTAGTCTGCTGCAATATGATGATGATGATGAT 480  
DB 421 AAAAGAAACCAAAAGAGGTGTAGTCTGCTGCAATATGATGATGATGATGATGAT 480  
QY 481 TACGAAATGATTAACCAACAACAGATGCTTCTTAATATGAAACGAAATTAACGGT 540  
DB 481 TACGAAATGATTAACCAACAACAGATGCTTCTTAATATGAAACGAAATTAACGGT 540  
QY 541 GTTCATATTTATCTTAAATAATGTTAGCAATGATGATGATGATGATGATGATGAT 600  
DB 541 GTTCATATTTATCTTAAATAATGTTAGCAATGATGATGATGATGATGATGATGAT 600  
QY 601 GGAACCTGTAAGAAATGAGATTAATGAGCAATTTGTTATTTCTAAAGCGAAGGC 660  
DB 601 GGAACCTGTAAGAAATGAGATTAATGAGCAATTTGTTATTTCTAAAGCGAAGGC 660  
QY 661 TCACGAGCAAGTAAATATATCCAGAGATCAAGATGATTAATATATATGAGCAAG 720  
DB 661 TCACGAGCAAGTAAATATATCCAGAGATCAAGATGATTAATATATATGAGCAAG 720  
QY 721 GATTAAGAAACGCAAAACGCTTTATTTACTGAGAAAGATTAATGATGATGATGAT 780  
DB 721 GATTAAGAAACGCAAAACGCTTTATTTACTGAGAAAGATTAATGATGATGATGAT 780  
QY 781 TTCAAGAAAGCAAGATGAAACGAGATTAATCAATTAATGAGATTAATGAGATTAAT 840  
DB 781 TTCAAGAAAGCAAGATGAAACGAGATTAATCAATTAATGAGATTAATGAGATTAAT 840  
QY 841 TATATTTTGAAGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAAT 900  
DB 841 TATATTTTGAAGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAAT 900  
QY 901 ACACATTTTACATTTGAGCAACATCAATCAACCTGTTGAAATGAAATGAAATGAT 960  
DB 901 ACACATTTTACATTTGAGCAACATCAATCAACCTGTTGAAATGAAATGAAATGAT 960  
QY 961 ACCTGTAAGTGAATTAACCAACCAAGCTTGAATGTAATGATGATGATGATGAT 1020  
DB 961 ACCTGTAAGTGAATTAACCAACCAAGCTTGAATGTAATGATGATGATGATGAT 1020

Db 955 ACCTCTAAAGTTGATTAACACACCAAGCTTAGATGTGAAAGTGGCAATTGGCCGA 1014  
Qy 1021 AAAATTAATATCAAAATTTCTGTAATATTTCCATTGGGGATTGGCAGCAAAAGAGCCGAC 1080  
Db 1015 AAAATTAATATCAAAATTTCTGTAATATTTCCATTGGGGATTGGCAGCAAAAGAGCCGAC 1074  
Qy 1081 GCTAATTAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140  
Db 1075 GCTAATTAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1134  
Qy 1141 AACGTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1200  
Db 1135 AACGTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1194  
Qy 1201 GAAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1260  
Db 1195 GAAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1254  
Qy 1261 ATTCTAGCTTAACACACGAGCGGACACCTAAATTCGTTTACTTTATGATGATGATGATGAT 1320  
Db 1255 ATTCTAGCTTAACACACGAGCGGACACCTAAATTCGTTTACTTTATGATGATGATGATGAT 1314  
Qy 1321 AAAGCAGATCTTACGAAAGGCTTTAAATTAAGAGCGAATGATTAACGATCATACCGAC 1380  
Db 1315 AAAGCAGATCTTACGAAAGGCTTTAAATTAAGAGCGAATGATTAACGATCATACCGAC 1374  
Qy 1381 GACCAACACCAACCACTGTTGAAGTTGATGATGATGATGATGATGATGATGATGATGATGAT 1440  
Db 1375 GACCAACACCAACCACTGTTGAAGTTGATGATGATGATGATGATGATGATGATGATGATGAT 1434  
Qy 1441 GCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1500  
Db 1435 GCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1494  
Qy 1501 AGCGACACGCAAAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1560  
Db 1495 AGCGACACGCAAAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1554  
Qy 1561 ACAAAGCTGAAGCACTTCTTTTACACACGCGTGAATGATGATGATGATGATGATGATGATGAT 1620  
Db 1555 ACAAAGCTGAAGCACTTCTTTTACACACGCGTGAATGATGATGATGATGATGATGATGATGAT 1614  
Qy 1621 CTTAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1680  
Db 1615 CTTAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1674  
Qy 1681 ACAAATCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1740  
Db 1675 ACAAATCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1734  
Qy 1741 TCACGAGAAAGTACCAACCAACCAAGGATGATGATGATGATGATGATGATGATGATGATGAT 1800  
Db 1735 TCACGAGAAAGTACCAACCAACCAAGGATGATGATGATGATGATGATGATGATGATGATGAT 1794  
Qy 1801 ATCTAGCTTACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1860  
Db 1795 ATCTAGCTTACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1854  
Qy 1861 AGACGTAGAAAAAAGAAATGCT 1881  
Db 1855 AGACGTAGAAAAAAGAAATGCT 1875

## RESULT 10

ADV16333 ID ADV16333 standard; DNA; 1875 BP.

XX AC ADV16333;

XX XX 24-FEB-2005 (first entry)

XX DE E. faecalis V583 hyperimmune serum reactive antigen DNA - SEQ ID 28.

KW antigen; antibacterial; vaccine; enterococcus infection; infection;  
KW pharyngitis; impetigo; rheumatic fever; antipyretic; antirheumatic;  
KW immunosuppressive; sepsis; ds; gene.  
OS Enterococcus faecalis V583.  
XX  
PN MO2004106367-A2.  
XX  
PD 09-DEC-2004.  
XX  
XX 26-MAY-2004; 2004WO-BP005664.  
XX  
XX 30-MAY-2003; 2003BP-00450137.  
XX  
XX (INRB-) INTERCELL AG.  
PI Meinke A, Nagy B, Hanner M, Gelbmann D;  
XX  
XX MPI; 2005-039707/04.  
DR P-PSDB; ADV16503.  
XX  
XX Novel isolated nucleic acid molecule encoding hyperimmune serum reactive  
PT antigen e.g., EP0020, EP0032, EP0062, EP0149, EP0196, EP0253, EP0270,  
PT EP0298 or its fragment, useful for producing vaccine against enterococcal  
PT infection.  
XX  
XX Example 3; SEQ ID NO 28; 175bp; English.  
XX  
XX The invention relates to a novel isolated nucleic acid molecule encoding  
CC a hyperimmune serum reactive antigen e.g. SEQ ID 171-340, 357-372 or 425-  
CC 476, or its fragment. The antigen of the invention demonstrates  
CC antibacterial activities and may be useful for producing a pharmaceutical  
CC preparation, particularly a vaccine, against enterococcal infection.  
CC Conditions associated with bacterial infection which may be prevented or  
CC treated include bacterial pharyngitis, scarlet fever, impetigo, rheumatic  
CC fever, necrotizing fasciitis and sepsis in humans. The current sequence  
CC is that of an Enterococcus faecalis V583 hyperimmune serum reactive  
CC antigen DNA of the invention.  
SQ Sequence 1875 BP; 691 A; 317 C; 400 G; 467 T; 0 U; 0 Other;  
XX  
XX Query Match 98.8%; Score 1859; DB 14; Length 1875;  
XX Best Local Similarity 99.7%; Pred. No. 0;  
XX Matches 1875; Conservative 0; Mismatches 0; Indels 6; Gaps 1;  
Qy 1 ATGAAGCAATTAATAAAGTTGGTACACCGTTAGTCTGTACTAATTTTGGCATT 60  
Db 1 ATGAAGCAATTAATAAAGTTGGTACACCGTTAGTCTGTACTAATTTTGGCATT 60  
Qy 61 TTCACAGCTGATTTGGGACCAACCTGATTTGAGAAAGAAATGGGAGGCGACAG 120  
Db 61 TTCACAGCTGATTTGGGACCAACCTGATTTGAGAAAGAAATGGGAGGCGACAG 120  
Qy 121 CTCGTGATTTCAAAAAGAAATGACGATTTTACAGATCCGTTATTTCAAAATAGCGGG 180  
Db 121 CTCGTGATTTCAAAAAGAAATGACGATTTTACAGATCCGTTATTTCAAAATAGCGGG 180  
Qy 181 AAAGAAATGAGCGATTTGATTAATATCAAGAGCTGGCAGATGATGATGATGATGATGAT 240  
Db 181 AAAGAAATGAGCGATTTGATTAATATCAAGAGCTGGCAGATGATGATGATGATGATGAT 240  
Qy 241 AACGTGAGCAAGCAATTTTACAGAGACGAGCGGAGCGCAAGGTTGATGATGATGATGAT 300  
Db 241 AACGTGAGCAAGCAATTTTACAGAGACGAGCGGAGCGCAAGGTTGATGATGATGATGAT 300  
Qy 301 CAAGCTGTCAAAGTTTAACTCTGGGAAACCTGTGTCTCAAGGAAACCAACCGATCAAT 360  
Db 301 CAAGCTGTCAAAGTTTAACTCTGGGAAACCTGTGTCTCAAGGAAACCAACCGATCAAT 360  
Qy 361 GGGATGTCACTGTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 420  
Db 361 GGGATGTCACTGTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 420



```
QY 421 AAAAGAAACCAAAAGAGGTGTAAGTCTGCTACGAATATGATGGTGCGTTCCAGTT 480
    |||
Db 421 AAAAGAAACCAAAAGAGGTGTAAGTCTGCTACGAATATGATGGTGCGTTCCAGTT 480
QY 481 TACGAATGATCAGCAAAACAGATGGTTCCTATTAATATGAAACAGAAATTAAGCGTT 540
    |||
Db 481 TACGAATGATCAGCAAAACAGATGGTTCCTATTAATATGAAACAGAAATTAAGCGTT 540
QY 541 GTTCATATTTATCCTTAATAATGTGGTAGCCATGATGATGTTTACATGTGAAAAAGTA 600
    |||
Db 541 GTTCATATTTATCCTTAATAATGTGGTAGCCATGATGATGTTTACATGTGAAAAAGTA 600
QY 601 GGAACCTGCTGAATAATGAAAGATTAAATGCGCGAAGATTTGTTATTTCTAAAGCGAAGC 660
    |||
Db 601 GGAACCTGCTGAATAATGAAAGATTAAATGCGCGAAGATTTGTTATTTCTAAAGCGAAGC 660
QY 655 GGAACCTGCTGAATAATGAAAGATTAAATGCGCGAAGATTTGTTATTTCTAAAGCGAAGC 654
    |||
Db 655 GGAACCTGCTGAATAATGAAAGATTAAATGCGCGAAGATTTGTTATTTCTAAAGCGAAGC 654
QY 661 TCACCAAGGCAAGTAAATATATATCCAAAGAGTCAAGATGATTTATATCATGACAAAG 720
    |||
Db 661 TCACCAAGGCAAGTAAATATATATCCAAAGAGTCAAGATGATTTATATCATGACAAAG 720
QY 655 TCACCAAGGCAAGTAAATATATATCCAAAGAGTCAAGATGATTTATATCATGACAAAG 714
    |||
Db 655 TCACCAAGGCAAGTAAATATATATCCAAAGAGTCAAGATGATTTATATCATGACAAAG 714
QY 721 GATTAAGAAACAAAGCAAAAGCTTTATATCTGGGAAAAAGTTATGAATTTGGCAAAATGAT 780
    |||
Db 721 GATTAAGAAACAAAGCAAAAGCTTTATATCTGGGAAAAAGTTATGAATTTGGCAAAATGAT 780
QY 715 GATTAAGAAACAAAGCAAAAGCTTTATATCTGGGAAAAAGTTATGAATTTGGCAAAATGAT 774
    |||
Db 715 GATTAAGAAACAAAGCAAAAGCTTTATATCTGGGAAAAAGTTATGAATTTGGCAAAATGAT 774
QY 781 TTCACAGAAAGCAGAAATGAAACGGGAGAAATTAAACAGTTAAAAATCTTGAGGTGGTTG 840
    |||
Db 781 TTCACAGAAAGCAGAAATGAAACGGGAGAAATTAAACAGTTAAAAATCTTGAGGTGGTTG 840
QY 775 TTCACAGAAAGCAGAAATGAAACGGGAGAAATTAAACAGTTAAAAATCTTGAGGTGGTTG 834
    |||
Db 775 TTCACAGAAAGCAGAAATGAAACGGGAGAAATTAAACAGTTAAAAATCTTGAGGTGGTTG 834
QY 841 TATATTTTGAAGAAAGTAAAGCTCCAAATTAATGCAAAATTAATTGAAAAATCAACAAA 900
    |||
Db 841 TATATTTTGAAGAAAGTAAAGCTCCAAATTAATGCAAAATTAATTGAAAAATCAACAAA 900
QY 835 TATATTTTGAAGAAAGTAAAGCTCCAAATTAATGCAAAATTAATTGAAAAATCAACAAA 894
    |||
Db 835 TATATTTTGAAGAAAGTAAAGCTCCAAATTAATGCAAAATTAATTGAAAAATCAACAAA 894
QY 901 ACACATTTTACAAATGGAAGCAAAACATCAACACCTGTGAAAAAAACATCAAAATGAT 960
    |||
Db 901 ACACATTTTACAAATGGAAGCAAAACATCAACACCTGTGAAAAAAACATCAAAATGAT 960
QY 895 ACACATTTTACAAATGGAAGCAAAACATCAACACCTGTGAAAAAAACATCAAAATGAT 954
    |||
Db 895 ACACATTTTACAAATGGAAGCAAAACATCAACACCTGTGAAAAAAACATCAAAATGAT 954
QY 961 ACCCTTAAAGTTGATTAACCAACACCAAGCTTAAAGTGAAGTGGCAATTGGCGAA 1020
    |||
Db 961 ACCCTTAAAGTTGATTAACCAACACCAAGCTTAAAGTGAAGTGGCAATTGGCGAA 1020
QY 955 ACCCTTAAAGTTGATTAACCAACCAACAGCTTAAGTGAAGTGGCAATTGGCGAA 1014
    |||
Db 955 ACCCTTAAAGTTGATTAACCAACCAACAGCTTAAGTGAAGTGGCAATTGGCGAA 1014
QY 1021 AAAATTTAATATCAAAATTTCTGTAAATATTCATTTGGGATTTGCAGACAAAGAGCGAC 1080
    |||
Db 1021 AAAATTTAATATCAAAATTTCTGTAAATATTCATTTGGGATTTGCAGACAAAGAGCGAC 1080
QY 1015 AAAATTTAATATCAAAATTTCTGTAAATATTCATTTGGGATTTGCAGACAAAGAGCGAC 1074
    |||
Db 1015 AAAATTTAATATCAAAATTTCTGTAAATATTCATTTGGGATTTGCAGACAAAGAGCGAC 1074
QY 1081 GCTATTAATATAGTCAAAATTCATTTAGTGAATTAACAGATGCGACCTTAATCTTTGAT 1140
    |||
Db 1081 GCTATTAATATAGTCAAAATTCATTTAGTGAATTAACAGATGCGACCTTAATCTTTGAT 1140
QY 1075 GCTATTAATATAGTCAAAATTCATTTAGTGAATTAACAGATGCGACCTTAATCTTTGAT 1134
    |||
Db 1075 GCTATTAATATAGTCAAAATTCATTTAGTGAATTAACAGATGCGACCTTAATCTTTGAT 1134
QY 1141 AACGTGACTTCTGAGAGATATGCTTAATGCTTAATATGATGGGATTAACAGTATGCTCCT 1200
    |||
Db 1141 AACGTGACTTCTGAGAGATATGCTTAATGCTTAATATGATGGGATTAACAGTATGCTCCT 1200
QY 1135 AACGTGACTTCTGAGAGATATGCTTAATGCTTAATATGATGGGATTAACAGTATGCTCCT 1194
    |||
Db 1135 AACGTGACTTCTGAGAGATATGCTTAATGCTTAATATGATGGGATTAACAGTATGCTCCT 1194
QY 1201 GAAAATTTATCAAGTGAATGCAACAGCAATGGCTTCACTGTGCGCTTAATCCAGCTAT 1260
    |||
Db 1201 GAAAATTTATCAAGTGAATGCAACAGCAATGGCTTCACTGTGCGCTTAATCCAGCTAT 1260
QY 1195 GAAAATTTATCAAGTGAATGCAACAGCAATGGCTTCACTGTGCGCTTAATCCAGCTAT 1254
    |||
Db 1195 GAAAATTTATCAAGTGAATGCAACAGCAATGGCTTCACTGTGCGCTTAATCCAGCTAT 1254
QY 1261 ATTCTAGCGCTAAACAGGCGGCAACATCAAAATTCGTTTACTTAATGATTTAAATGAA 1320
    |||
Db 1261 ATTCTAGCGCTAAACAGGCGGCAACATCAAAATTCGTTTACTTAATGATTTAAATGAA 1320
QY 1255 ATTCTAGCGCTAAACAGGCGGCAACATCAAAATTCGTTTACTTAATGATTTAAATGAA 1314
    |||
Db 1255 ATTCTAGCGCTAAACAGGCGGCAACATCAAAATTCGTTTACTTAATGATTTAAATGAA 1314
QY 1331 AAAAGAGATCCTACGAAAGGCTTTAAAAATGAGCGCAATGTTGATTAACGCTCATACGAC 1380
    |||
Db 1331 AAAAGAGATCCTACGAAAGGCTTTAAAAATGAGCGCAATGTTGATTAACGCTCATACGAC 1380
QY 1315 AAAAGAGATCCTACGAAAGGCTTTAAAAATGAGCGCAATGTTGATTAACGCTCATACGAC 1374
    |||
Db 1315 AAAAGAGATCCTACGAAAGGCTTTAAAAATGAGCGCAATGTTGATTAACGCTCATACGAC 1374
QY 1381 GACCAAAACCAACCAACTGTTGAAGTTGCAAGGTGGGAAACGTTTCACTTAAGTCGAT 1440
    |||
Db 1381 GACCAAAACCAACCAACTGTTGAAGTTGCAAGGTGGGAAACGTTTCACTTAAGTCGAT 1440
QY 1375 GACCAAAACCAACCAACTGTTGAAGTTGCAAGGTGGGAAACGTTTCACTTAAGTCGAT 1434
    |||
Db 1375 GACCAAAACCAACCAACTGTTGAAGTTGCAAGGTGGGAAACGTTTCACTTAAGTCGAT 1434
QY 1441 GCGGATGTAAGCAGCAACAGCTTTGGGGAGGCTTCTTTTGGTCCGTGATCAAAAC 1500
    |||
Db 1441 GCGGATGTAAGCAGCAACAGCTTTGGGGAGGCTTCTTTTGGTCCGTGATCAAAAC 1500
QY 1435 GCGGATGTAAGCAGCAACAGCTTTGGGGAGGCTTCTTTTGGTCCGTGATCAAAAC 1494
    |||
Db 1435 GCGGATGTAAGCAGCAACAGCTTTGGGGAGGCTTCTTTTGGTCCGTGATCAAAAC 1494
QY 1501 AGCGACACAGCAAAATTAATTTGAAAAATCGATGAAACAAACGAAAGCAACCTTGCGTGAAA 1560
    |||
Db 1501 AGCGACACAGCAAAATTAATTTGAAAAATCGATGAAACAAACGAAAGCAACCTTGCGTGAAA 1560
    |||
```

```
Db 1495 AGCGACACAGCAAAATTTTGAATAATCGATGAAACAAAGAAAGCAACCTTGGGTGAAA 1554
    |||
QY 1561 ACAAAAGCTGAAGCAACTACTTTTACAACAACGCGTGAATGATTAATGATATCAACAGG 1620
    |||
Db 1555 ACAAAAGCTGAAGCAACTACTTTTACAACAACGCGTGAATGATTAATGATATCAACAGG 1614
    |||
QY 1621 CTTAAATCGGTAACCTATTAATTAAGAAGAACTGATAGCTCCGATGATTAATGCTTTGTA 1680
    |||
Db 1615 CTTAAATCGGTAACCTATTAATTAAGAAGAACTGATAGCTCCGATGATTAATGCTTTGTA 1674
    |||
QY 1681 ACAAAATCGGATTTGAATTTGTGCTCAATGCAATCATATGCAACAAAGAAAACTAGTT 1740
    |||
Db 1675 ACAAAATCGGATTTGAATTTGTGCTCAATGCAATCATATGCAACAAAGAAAACTAGTT 1734
    |||
QY 1741 TCACCAAGAAAGATACCAACCAACAAAGATCCTTAACCTTCAACAGGTGGCAAGGA 1800
    |||
Db 1735 TCACCAAGAAAGATACCAACCAACAAAGATCCTTAACCTTCAACAGGTGGCAAGGA 1794
    |||
QY 1801 ATCTAGTTTACTTAAGAAAGTGGCGAGTCTGCTACTTAATGCAAGAGTCTACTTTGCT 1860
    |||
Db 1795 ATCTAGTTTACTTAAGAAAGTGGCGAGTCTGCTACTTAATGCAAGAGTCTACTTTGCT 1854
    |||
QY 1861 AGACGTAGAAAAAGAAATGCT 1881
    |||
Db 1855 AGACGTAGAAAAAGAAATGCT 1875
    |||

RESULT 11
AAAX20109
ID AAAX20109 standard; DNA; 1687 BP.
XX
AC AAAX20109;
XX
DT 20-APR-1999 (first entry)
XX
DE Enterococcus faecalis EF058 gene fragment.
XX
KW Enterococcus faecalis; infection; vaccine; immune response; diagnosis;
XX detection; attenuation; antigenic; ss.
XX
OS Enterococcus faecalis.
XX
PN WO980554-A2.
XX
PD 12-NOV-1998.
XX
PF 04-MAY-1998; 98WO-US008959.
XX
PR 06-MAY-1997; 97US-0044031P.
XX 16-MAY-1997; 97US-004655P.
XX 14-NOV-1997; 97US-0066009P.
XX
PA (HDMA-) HUMAN GENOME SCI INC.
XX
PI Kunsch CA, Choi GH, Bailey C, Hromockyj A;
XX
DR WPI: 1999-070095/06.
XX
DR P-PSDB; AAY00119.
XX
XX
PT New isolated Enterococcus faecalis polynucleotides - used to develop
XX products for the detection of Enterococcus and for use in vaccines for
XX prevention or attenuation of Enterococcus infection.
XX
PS Claim 1; Page 134; 301pp; English.
XX
XX
CC The present sequence encodes an antigenic polypeptide fragment isolated
XX from Enterococcus faecalis. The present invention describes genes,
XX proteins and antigenic polypeptides isolated from E. faecalis. The
XX proteins can be used in vaccines for preventing or attenuating an
XX infection caused by a member of the Enterococcus genus in an animal. They
XX can also be used for detecting Enterococcus antibodies in a sample. The
XX nucleotide sequences can be used for detecting Enterococcus nucleic
```



CC acids. Products from the present invention can also be used for screening  
CC compounds to identify agonists and antagonists of *E. faecalis* protein  
CC activity

XX Sequence 1687 BP; 632 A; 284 C; 360 G; 411 T; 0 U; 0 Other;

Query Match 89.7%; Score 1687; DB 2; Length 1687;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1687; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 96 AGAAGAAAATGGGAGAGCGCAACGCTGTGATTCACAAAAGAAAATGACGATTTAC 155
DB 1 AGAAGAAAATGGGAGAGCGCACAGCTGTGATTCACAAAAGAAAATGACGATTTAC 60
QY 156 AGATCCGCTTATTTCAAAATAGCGGAAAAGAAATGACGATTTGATTAATTCAGAGACT 215
DB 61 AGATCCGCTTATTTCAAAATAGCGGAAAAGAAATGACGATTTGATTAATTCAGAGACT 120
QY 216 GGCAGATGTGACGTTAGTATTTATATACGTGACGAAAGCAATTTTACGACAGACGCGC 275
DB 121 GGCAGATGTGACGTTAGTATTTATATACGTGACGAAAGCAATTTTACGACAGACGCGC 180
QY 276 AGCGCAAGCGTTGATGACGCTTAAACAGCTGTCCAAAGTTTAACTCTGGGAAACTGT 335
DB 181 AGCGCAAGCGTTGATGACGCTTAAACAGCTGTCCAAAGTTTAACTCTGGGAAACTGT 240
QY 336 TGCTCAAGAAACACCGATGCAATGCGAATGTCACTGTTCAGTTAATCTTAAAAACAAA 395
DB 241 TGCTCAAGAAACACCGATGCAATGCGAATGTCACTGTTCAGTTAATCTTAAAAACAAA 300
QY 396 TGGTAAAGATGACGATGTATACCATTAAGAAAGCAAAAGAGGGGTGATGTTGCTGTAC 455
DB 301 TGGTAAAGATGACGATGTATACCATTAAGAAAGCAAAAGAGGGGTGATGTTGCTGTAC 360
QY 456 GAATATGTGTGTGGCGTTCCAGTTTACGAAATGATCAAGCAACAGATGTTCTTATTA 515
DB 361 GAATATGTGTGTGGCGTTCCAGTTTACGAAATGATCAAGCAACAGATGTTCTTATTA 420
QY 516 ATATGGAACAGAAAGATAGCGGTGTTCTATTTATCTTAAAAATGTGTGACCAATGA 575
DB 421 ATATGGAACAGAAAGATAGCGGTGTTCTATTTATCTTAAAAATGTGTGACCAATGA 480
QY 576 TGGTATGTTACATGGAAGAAAAGTAGAACTGCTGAAGATGAAGATTAAATGGCGCAGA 635
DB 481 TGGTATGTTACATGGAAGAAAAGTAGAACTGCTGAAGATGAAGATTAAATGGCGCAGA 540
QY 636 ATTTGTTATTTCTAAGAGCGAAGGCTCAACAGGCAAGTAAATATATCCAGAGAGTCA 695
DB 541 ATTTGTTATTTCTAAGAGCGAAGGCTCAACAGGCAAGTAAATATATCCAGAGAGTCA 600
QY 696 AGATGATTAATATACATGACACGATTAAGAAAGACAGCAAAACGCTTTATTACTGGGA 755
DB 601 AGATGATTAATATACATGACACGATTAAGAAAGACAGCAAAACGCTTTATTACTGGGA 660
QY 756 AGTTATGAATTTGGCGAAAATGATTTCAAGAGCAGAGATGGAACGGGAGAAATTAC 815
DB 661 AGTTATGAATTTGGCGAAAATGATTTCAAGAGCAGAGATGGAACGGGAGAAATTAC 720
QY 816 AGTTAAAAATCTTGAGTTGGTTCGTATATTTTGAAGAAAGTAAAGCTCCAAATATATGC 875
DB 721 AGTTAAAAATCTTGAGTTGGTTCGTATATTTTGAAGAAAGTAAAGCTCCAAATATATGC 780
QY 876 AGAATTATTTAAAAATCAAAACAAACCATTTTCAATTTGAAGCAATCAACACCC 935
DB 781 AGAATTATTTAAAAATCAAAACAAACCATTTTCAATTTGAAGCAATCAACACCC 840
QY 936 TGTGAAAAAAACAGTCAAAAATGATACCTTAAAGTTAAAAACAACCAAGCTTGA 995
DB 841 TGTGAAAAAAACAGTCAAAAATGATACCTTAAAGTTAAAAACAACCAAGCTTGA 900
QY 996 TGGTAAAGATGTGCAATTTGGCAAAAAATTAATATCAAAATTTCTGTAATATTCATT 1055
DB 901 TGGTAAAGATGTGCAATTTGGCAAAAAATTAATATCAAAATTTCTGTAATATTCATT 960
```

```
QY 1056 GGGATTTGACAGCAAAAGAGCGACGCTTAATTAATACGTAATCAATTTAGTGAATA 1115
DB 961 GGGATTTGACAGCAAAAGAGCGACGCTTAATTAATACGTAATCAATTTAGTGAATA 1020
QY 1116 ACATGATGACGCTTAATCTTTTGAATACGTGACTCTGGAGAGTANGCTTATGCGTTATA 1175
DB 1021 ACATGATGACGCTTAATCTTTTGAATACGTGACTCTGGAGAGTANGCTTATGCGTTATA 1080
QY 1176 TGATGGGAAATACGATGATTTGCTCTGAAATTTATCAAGTGAACAGCAAAATGCGTT 1235
DB 1081 TGATGGGAAATACGATGATTTGCTCTGAAATTTATCAAGTGAACAGCAAAATGCGTT 1140
QY 1236 CACTGTCCCGCTTATTCAGCGTATATTTCTTACGTTAACACAGCGCGCACCTAAAT 1295
DB 1141 CACTGTCCCGCTTATTCAGCGTATATTTCTTACGTTAACACAGCGCGCACCTAAAT 1200
QY 1296 CGTTTACTTTATGACATTTAAATGAAGAAAAGCAGATCTTACGAAAGGCTTTAAAAATGAGGC 1355
DB 1201 CGTTTACTTTATGACATTTAAATGAAGAAAAGCAGATCTTACGAAAGGCTTTAAAAATGAGGC 1260
QY 1356 GAATGTTGATTAACGTCATACGACGACCAACACACCAACTGTTGAAGTTGTGACAG 1415
DB 1261 GAATGTTGATTAACGTCATACGACGACCAACACCAACTGTTGAAGTTGTGACAG 1320
QY 1416 TGGGAAAGCTTTCAATTAAGTCGATGCGGATGTGACACGACACAAAGCTTGGCGGAGC 1475
DB 1321 TGGGAAAGCTTTCAATTAAGTCGATGCGGATGTGACACGACACAAAGCTTGGCGGAGC 1380
QY 1476 TTCCCTTGTGCTGCGTGATCAAAAACAGGCAACAGCAAAATTTTGAATTCATGATAAC 1535
DB 1381 TTCCCTTGTGCTGCGTGATCAAAAACAGGCAACAGCAAAATTTTGAATTCATGATAAC 1440
QY 1536 AACGAAAGCAGCAACTTGGGTGAAAACAAAAGCTGAACCACTATTTTACAAACAGGC 1595
DB 1441 AACGAAAGCAGCAACTTGGGTGAAAACAAAAGCTGAACCACTATTTTACAAACAGGC 1500
QY 1596 TGATGGAATTAATGATATCAACAGGCTTAATACGCTATCTTATTTTGAAGAAATGCT 1655
DB 1501 TGATGGAATTAATGATATCAACAGGCTTAATACGCTATCTTATTTTGAAGAAATGCT 1560
QY 1656 AGCTCCGATGATTAATGCTTGTTPAACAAATCGAATGAATTTGTGCTCAATGAACAATC 1715
DB 1561 AGCTCCGATGATTAATGCTTGTTPAACAAATCGAATGAATTTGTGCTCAATGAACAATC 1620
QY 1716 ATATGGAACAAACAGAAAACCTAGTTTCAACGAAAAGATACCAACAAACAGAGTAC 1775
DB 1621 ATATGGAACAAACAGAAAACCTAGTTTCAACGAAAAGATACCAACAAACAGAGTAC 1680
QY 1776 CTTTACT 1782
DB 1681 CTTTACT 1687
```

RESULT 12  
ID ABN98094 standard; DNA; 1687 BP.  
XX  
AC ABN98094;  
DT 05-AUG-2002 (first entry)  
XX  
DE *E. faecalis* EF058 gene fragment.  
XX  
KW Enterococcus; vaccine; gastrointestinal disease; diagnosis; antibiotic;  
FN gene; *del.*  
OS Enterococcus faecalis.  
XX  
FN US2002045737-A1.  
XX  
PD 18-APR-2002.  
XX

PF 04-MAY-1998; 98US-00071035.  
XX 04-MAY-1998; 98US-00071035.  
XX (HUMA-) HUMAN GENOME SCI INC.  
PA Chol GH, Bailey C, Hromockyj A, Kunsch CA;  
XX WPI; 2002-425450/45.  
DR P-PSDB; ABP43338.  
XX  
XX New genes and polypeptides from *Enterococcus faecalis*, useful as vaccines  
PT for preventing, treating or attenuating an infection caused by a member  
PT of the *Enterococcus* genus in an animal, particularly *E. faecalis*.  
XX  
XX Claim 1; Page 93-94; 255pp; English.  
XX  
XX The present invention provides the protein and coding sequences of a  
CC number of polypeptides from *Enterococcus faecalis*. The proteins can be  
CC used as vaccines for preventing or attenuating an infection caused by a  
CC member of the *Enterococcus* genus in an animal, particularly *E. faecalis*.  
CC The polynucleotide is also useful for preventing or treating *E. faecalis*  
CC infection. The present sequence is a coding sequence of the invention  
XX  
SQ Sequence 1687 BP; 632 A; 284 C; 360 G; 411 T; 0 U; 0 Other;  
Query Match 89.7%; Score 1687; DB 6; Length 1687;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1687; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 96 AGAAGAAATGGGGAGAGCGACAGCTGATTCACAAAAGAAAATGACGGATTAC 155  
DB 1 AGAAGAAATGGGGAGAGCGACAGCTGATTCACAAAAGAAAATGACGGATTAC 60  
QY 156 AGATCGCTTATTCAAAATAGCGGAGAAAGAAATGACGGATTACCAAGACT 215  
DB 61 AGATCGCTTATTCAAAATAGCGGAGAAAGAAATGACGGATTACCAAGACT 120  
QY 216 GGCAGATGACGCTTATTCATTTATACGTGACGAAGAAATTTTACGACACGAGCGG 275  
DB 121 GGCAGATGACGCTTATTCATTTATACGTGACGAAGAAATTTTACGACACGAGCGG 180  
QY 276 AGGCGCAAGCGTTGATGACGCTTAAACAAGCTGCTCAAGTTTACTCTGGGAAACCTGT 335  
DB 181 AGGCGCAAGCGTTGATGACGCTTAAACAAGCTGCTCAAGTTTACTCTGGGAAACCTGT 240  
QY 336 TGCTCAGGAACCCAGCGATGCAATGGGAATGTCACTGTTCACTTAACCTTAAACAA 395  
DB 241 TGCTCAGGAACCCAGCGATGCAATGGGAATGTCACTGTTCACTTAACCTTAAACAA 300  
QY 396 TGGTAAAGATGACGCTTATTCATTTATACGTGACGAAGAAATTTTACGACACGAG 455  
DB 301 TGGTAAAGATGACGCTTATTCATTTATACGTGACGAAGAAATTTTACGACACGAG 360  
QY 456 GAATATGATGCTGCTGCCAGTTTACGAAATGATCAAGCAACAGATGCTTCTATTA 515  
DB 361 GAATATGATGCTGCTGCCAGTTTACGAAATGATCAAGCAACAGATGCTTCTATTA 420  
QY 516 ATATGGAACAGAAAGATGACGCTTATTTATCTTAAATTTGTTGACCAATGA 575  
DB 421 ATATGGAACAGAAAGATGACGCTTATTTATCTTAAATTTGTTGACCAATGA 480  
QY 576 TGGTAAAGATGACGCTTATTCATTTATACGTGACGAAGAAATTTTACGACACGAG 635  
DB 481 TGGTAAAGATGACGCTTATTCATTTATACGTGACGAAGAAATTTTACGACACGAG 540  
QY 636 ATTTGTTATTTCTTAAAGAGGCTCAACAGCAACAGTAAATATATCAAGAGTCA 695  
DB 541 ATTTGTTATTTCTTAAAGAGGCTCAACAGCAACAGTAAATATATCAAGAGTCA 600  
QY 696 AGATGATTTATTCATTCAGCAACGATTAAGAAACAGTAAATCTTTATTTACTGGGA 755  
DB 601 AGATGATTTATTCATTCAGCAACGATTAAGAAACAGTAAATCTTTATTTACTGGGA 660

QY 756 AAGTTATGAATTTGGCGGAAATGATTTCAAGAGCAGAGATGGAACGGGAAATTAAC 815  
DB 661 AAGTTATGAATTTGGCGGAAATGATTTCAAGAGCAGAGATGGAACGGGAAATTAAC 720  
QY 816 AGTTAAATATCTTGAGTTGGTTCGTATATTTTAAAGAAAGTAAAGTCCAAATATGC 875  
DB 721 AGTTAAATATCTTGAGTTGGTTCGTATATTTTAAAGAAAGTAAAGTCCAAATATGC 780  
QY 876 AGAATTAATTTGAATCAAAACAAACCAATTTCAATTTGAAGCAACATCAACACC 935  
DB 781 AGAATTAATTTGAATCAAAACAAACCAATTTCAATTTGAAGCAACATCAACACC 840  
QY 936 TGTGAAAAACAGCAAAAAATGATACCTTAAAGTTGATTAACCAACCAAGCTTGA 995  
DB 841 TGTGAAAAACAGCAAAAAATGATACCTTAAAGTTGATTAACCAACCAAGCTTGA 900  
QY 996 TGTGAAAGATGCGCAATTTGGCGGAAATTAATTAATCAAAATTTCTGTAATATTCAT 1055  
DB 901 TGTGAAAGATGCGCAATTTGGCGGAAATTAATTAATCAAAATTTCTGTAATATTCAT 960  
QY 1056 GGGATTTGACAAAGAAAGGCGAGCTTAATTAATACGTCAAATTCATTTAGTGAAT 1115  
DB 961 GGGATTTGACAAAGAAAGGCGAGCTTAATTAATACGTCAAATTCATTTAGTGAAT 1020  
QY 1116 ACATGATCAGCCTTAACTTTTATGATGATGACGCTTGAAGATGCTTATGCTTATTA 1175  
DB 1021 ACATGATCAGCCTTAACTTTTATGATGATGACGCTTGAAGATGCTTATGCTTATTA 1080  
QY 1176 TGATGGGATACAGTATGCTCTGAAATTAATTAATCAAGTGAACCAAGCAATGCTT 1235  
DB 1081 TGATGGGATACAGTATGCTCTGAAATTAATTAATCAAGTGAACCAAGCAATGCTT 1140  
QY 1236 CACGTGCGCGTTATTCAGGATTAATTCACGCTAACACCAAGGCGGACCTTAATTT 1295  
DB 1141 CACGTGCGCGTTATTCAGGATTAATTCACGCTAACACCAAGGCGGACCTTAATTT 1200  
QY 1296 CGTTTACTTTATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1355  
DB 1201 CGTTTACTTTATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1260  
QY 1356 GAATGTTGATTAACGCTTATTCAGGATTAATTAATTAATTAATTAATTAATTAAT 1415  
DB 1261 GAATGTTGATTAACGCTTATTCAGGATTAATTAATTAATTAATTAATTAATTAAT 1320  
QY 1416 TGGGAAGCTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1475  
DB 1321 TGGGAAGCTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1380  
QY 1476 TTCTTTGCTGCTGATCAAAACAGGACACAGCAATTAATTAATTAATTAATTAAT 1535  
DB 1381 TTCTTTGCTGCTGATCAAAACAGGACACAGCAATTAATTAATTAATTAATTAAT 1440  
QY 1536 AACGAAGCAGCACTTGGGTTGAAACAAAGCTGAACACTTTTCAACACGCG 1595  
DB 1441 AACGAAGCAGCACTTGGGTTGAAACAAAGCTGAACACTTTTCAACACGCG 1500  
QY 1596 TGAATGATTAAGTGAATCAAGGCTTAAATTAAGGCTTAAATTAATTAATTAATTAAT 1655  
DB 1501 TGAATGATTAAGTGAATCAAGGCTTAAATTAAGGCTTAAATTAATTAATTAATTAAT 1560  
QY 1656 AGCTCTGATGATTAATGCTTGTTAACAATCGAATTAATTAATTAATTAATTAAT 1715  
DB 1561 AGCTCTGATGATTAATGCTTGTTAACAATCGAATTAATTAATTAATTAATTAAT 1620  
QY 1716 ATATGCAACACAGAAACCTTATTTCAACGAAACAAAGTGAAGTAC 1775  
DB 1621 ATATGCAACACAGAAACCTTATTTCAACGAAACAAAGTGAAGTAC 1680  
QY 1776 CTTACT 1782  
DB 1681 CTTACT 1687

RESULT 13  
ACA88058  
ID ACA88058 standard; DNA; 1687 BP.  
AC ACA88058;  
DT 07-JUL-2003 (first entry)  
XX E. faecalis novel gene #110.  
DE  
XX Gene; ds; endocarditis; bacteraemia; urinary tract infection; URI;  
KM intrabdominal infection; soft tissue infection; neonatal sepsis;  
KW vaccine.  
XX Enterococcus faecalis.  
OS  
XX US2003017495-A1.  
PN  
XX 23-JAN-2003.  
PD  
XX 29-JUL-2002; 2002US-00206576.  
PF  
XX 06-MAY-1997; 97US-0044031P.  
PR 16-MAY-1997; 97US-0046655P.  
PR 14-NOV-1997; 97US-0066009P.  
PR 04-MAY-1998; 98US-00071035.  
XX  
PA (HUMA-) HUMAN GENOME SCI INC.  
XX  
XX Choi GH, Bailey C, Hromockyj A, Kunsch CA;  
PI WPI; 2003-416890/39.  
DR P-PsDB; ABB88366.  
XX  
XX New nucleic acid molecules and polypeptides from Enterococcus faecalis,  
PT useful as vaccines for preventing or attenuating an enterococcal  
PT infection in an animal, or for identifying Enterococcus faecalis in  
PT biological samples.  
XX  
XX Claim 1; Page; 40pp; English.  
XX  
XX The invention relates to a new isolated nucleic acid molecule comprising  
CC a polynucleotide isolated from Enterococcus faecalis appearing as  
CC ACA87949-ACA88196 (or sequences complementary to them or 95% identical to  
CC them). Also included are the proteins encoded by the above nucleic acids,  
CC making a recombinant vector (comprising inserting the isolated nucleic  
CC acid molecule cited above into a vector), a host cell comprising the  
CC vector (used to produce the protein), an isolated antibody specific for  
CC the polypeptides, a hybridoma that produces the antibody, an isolated  
CC polypeptide antigen comprising an amino acid sequence of an Enterococcus  
CC faecalis epitope listed in the specification, a vaccine comprising one or  
CC more E. faecalis polypeptides (and a pharmaceutical diluent, carrier or  
CC excipient) where the polypeptide elicits protective antibodies in an  
CC animal to a member of the genus Enterococcus; preventing or attenuating  
CC an infection caused by a member of the genus Enterococcus in an animal  
CC comprising administering to the animal the polypeptide and detecting  
CC Enterococcus nucleic acids in a biological sample. The E. faecalis  
CC nucleic acid molecules and polypeptides are useful as vaccines for  
CC preventing or attenuating an enterococcal infection in an animal (e.g.  
CC infection, soft tissue infection and neonatal sepsis). The polypeptides  
CC are also useful for detecting Enterococcus aureus in immunosays, as  
CC epitope tags, as molecular weight markers, or for generating antibodies  
CC that specifically bind E. faecalis polypeptides. The nucleic acid  
CC molecules are also useful as probes for gene mapping, or for identifying  
CC E. faecalis in biological samples. The kit and methods are useful for  
CC detecting Enterococcus antibodies or nucleic acid molecules in a  
CC biological sample. The present sequence is a novel E. faecalis nucleic  
CC acid of the invention. Note: The sequence data for this patent did not  
CC form part of the printed specification, but was obtained in electronic  
CC format directly from the USPTO at  
CC seqdata.uspto.gov/sequence.html?docID=20030017495

XX  
SQ Sequence 1687 BP; 632 A; 284 C; 360 G; 411 T; 0 U; 0 Other;  
Query Match 89.7%; Score 1687; DB 8; Length 1687;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1687; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 96 AGAAGAAAATGGGAGAGGCGACAGCTGTGATTCACAAAAAGAAATGACGATTTACC 155  
DB 1 AGAAGAAAATGGGAGAGGCGACAGCTGTGATTCACAAAAAGAAATGACGATTTACC 60  
QY 156 AGATCGCTTATTCACAAATAGCGGGAAGAAATGAGCGAGTTGATTAATACAGACT 215  
DB 61 AGATCGCTTATTCACAAATAGCGGGAAGAAATGAGCGAGTTGATTAATACAGACT 120  
QY 216 GGCAGATGTGACGTTAGTATTTATPAACGTGACGAACGAATTTACGAGCAAGACGGC 275  
DB 121 GGCAGATGTGACGTTAGTATTTATPAACGTGACGAACGAATTTACGAGCAAGACGGC 180  
QY 276 AGCGCAAGCGCTTGATGACGCTTAACCAAGCTGTCCAAAGTTTAACCTCTGGGAAACCTGT 335  
DB 181 AGCGCAAGCGCTTGATGACGCTTAACCAAGCTGTCCAAAGTTTAACCTCTGGGAAACCTGT 240  
QY 336 TGCTCAAGGAACCAACCGATGCAATGGAATGTCACTGTTCACTTAACCTTAACAAACAAA 395  
DB 241 TGCTCAAGGAACCAACCGATGCAATGGAATGTCACTGTTCACTTAACCTTAACAAACAAA 300  
QY 396 TGGTAAGATGACGATGTATACCATTAAGAAAGAACCAAAAGAGGTGTATGTTGCTGTAC 455  
DB 301 TGGTAAGATGACGATGTATACCATTAAGAAAGAACCAAAAGAGGTGTATGTTGCTGTAC 360  
QY 456 GAATATGTTGTTGGGCTTCCGATTTACGAATGATCAAGCAACAGATGTTCTATTA 515  
DB 361 GAATATGTTGTTGGGCTTCCGATTTACGAATGATCAAGCAACAGATGTTCTATTA 420  
QY 516 ATATGAAACAGAAATAGCGGTTTCATTTATCTTAATAATGTTGTTGCTCAATGA 575  
DB 421 ATATGAAACAGAAATAGCGGTTTCATTTATCTTAATAATGTTGTTGCTCAATGA 480  
QY 576 TGGTATGTTACATGTGAAAAAAGTAGAAGCTGTGAAAAATGAAGATTAAATGGCGCAGA 635  
DB 481 TGGTATGTTACATGTGAAAAAAGTAGAAGCTGTGAAAAATGAAGATTAAATGGCGCAGA 540  
QY 636 ATTTGTTATTTCTTAACCGAAGGCTCAACGACGACGTAATATATTCAGAGGTCAA 695  
DB 541 ATTTGTTATTTCTTAACCGAAGGCTCAACGACGACGTAATATATTCAGAGGTCAA 600  
QY 696 AGATGATTTATATACATGAGCAACGATTAAGAAACAGCAAAAACGTTTATTAATGCGGAA 755  
DB 601 AGATGATTTATATACATGAGCAACGATTAAGAAACAGCAAAAACGTTTATTAATGCGGAA 660  
QY 756 AAGTTATGAAATTTGGCGAAATGATTTCAAGAAACAGAAATGGAACGGAGAAATTAAC 815  
DB 661 AAGTTATGAAATTTGGCGAAATGATTTCAAGAAACAGAAATGGAACGGAGAAATTAAC 720  
QY 816 AGTTAAATATCTGAGGTTGTTGCTATATTTTGAAGAAAGTTAAAGCTCCAAATTAATGC 875  
DB 721 AGTTAAATATCTGAGGTTGTTGCTATATTTTGAAGAAAGTTAAAGCTCCAAATTAATGC 780  
QY 876 AGAATTAATTAATCAACAAACAAACAAACAACTTTCATTTGAAGCAATCAACACCC 935  
DB 781 AGAATTAATTAATCAACAAACAAACAAACAACTTTCATTTGAAGCAATCAACACCC 840  
QY 936 TGGTGAAGAAACAGTCAAAATGATACCTGTAAAGTTGATTAACCAACCAAGCTTGA 995  
DB 841 TGGTGAAGAAACAGTCAAAATGATACCTGTAAAGTTGATTAACCAACCAAGCTTGA 900  
QY 996 TGGTGAAGAAATGCGAATGCGGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1055  
DB 901 TGGTGAAGAAATGCGAATGCGGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 960  
QY 1056 GGGGATTCAGACAAAGAAAGCGACGCTTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1115

Db 961 GGGGATGCGACGCAAGAGGCGGACGCTAATTAATACGTAATTAATCAATTAATGATTGATTA 1020  
QY 1116 ACATGATGACGCTTAATCTTTGATTAAGTGACTTCTGGAAGTATGCTTAATGCTTTATA 1175  
Db 1021 ACATGATGACGCTTAATCTTTGATTAAGTGACTTCTGGAAGTATGCTTAATGCTTTATA 1080  
QY 1176 TGATGGGATACAGATGATTGCTCTGAAAAATTATCAAGTGAAGTGAACAAGAAATGAGCTT 1235  
Db 1081 TGATGGGATACAGATGATTGCTCTGAAAAATTATCAAGTGAAGTGAACAAGAAATGAGCTT 1140  
QY 1236 CACTGTCCGCTTAATCCAGCGTATATTCCTACGCTAACACGAGCGGACACTAAAT 1295  
Db 1141 CACTGTCCGCTTAATCCAGCGTATATTCCTACGCTAACACGAGCGGACACTAAAT 1200  
QY 1296 CGTTACTTTTATGCTTTTAAATGAAAAAGAGTCTCTAGAAAGCTTTTAAATGAGCG 1355  
Db 1201 CGTTACTTTTATGCTTTTAAATGAAAAAGAGTCTCTAGAAAGCTTTTAAATGAGCG 1260  
QY 1356 GAATGTTGATTAAGGTCATACCGACGACCAACCAACCACTGTTGAAGTGTGACAG 1415  
Db 1261 GAATGTTGATTAAGGTCATACCGACGACCAACCAACCACTGTTGAAGTGTGACAG 1320  
QY 1416 TGGGAAAGCTTCAATTAAGTGTGATGCGATGTGACAGGACCAAGCTTGGCGGAGC 1475  
Db 1321 TGGGAAAGCTTCAATTAAGTGTGATGCGATGTGACAGGACCAAGCTTGGCGGAGC 1380  
QY 1476 TTCTTTGCTGCTGATCAAAACAGGACCAAGCAATTTTGAATGATGAAG 1535  
Db 1381 TTCTTTGCTGCTGATCAAAACAGGACCAAGCAATTTTGAATGATGAAG 1440  
QY 1536 AACGAAAGCAGCACTGGGTGAAAAACAAGCTGAGCACTGTTTACCAACAGCG 1595  
Db 1441 AACGAAAGCAGCACTGGGTGAAAAACAAGCTGAGCACTGTTTACCAACAGCG 1500  
QY 1596 TGATGATTAAGTGTATATCAAGGCTTAATACGCTTATTTTGAAGAACTGT 1655  
Db 1501 TGATGATTAAGTGTATATCAAGGCTTAATACGCTTATTTTGAAGAACTGT 1560  
QY 1656 AGCTCTGATGATTAATGCTTTTAAACAAGTGAATGATGATGAACAATC 1715  
Db 1561 AGCTCTGATGATTAATGCTTTTAAACAAGTGAATGATGATGAACAATC 1620  
QY 1716 ATATGCAACAACAGAAACCTAGTTTCAACAGAAAAAGTACCAACCAACAGAACTGT 1775  
Db 1621 ATATGCAACAACAGAAACCTAGTTTCAACAGAAAAAGTACCAACCAACAGAACTGT 1680  
QY 1776 CTTACTT 1782  
Db 1681 CTTACTT 1687

RESULT 14  
ABX61664  
ID ABX61664 standard; DNA; 1687 BP.  
XX  
AC ABX61664;  
XX  
DT 26-FEB-2003 (first entry)  
XX  
DE Enterococcus faecalis EF040 polynucleotide #110.  
XX  
KM EF040; gene; ds; immunostimulant; antibacterial; gene mapping.  
XX  
OS Enterococcus faecalis.  
XX  
FN US6448043-B1.  
XX  
PD 10-SEP-2002.  
XX  
PF 04-MAY-1998; 98US-00071035.  
XX  
PR 06-MAY-1997; 97US-0044031P.  
PR 16-MAY-1997; 97US-004655BP.

PR 14-NOV-1997; 97US-0066099P.  
PR 14-NOV-1997; 97US-0066099P.  
XX  
PA (HOMA-) HUMAN GENOME SCI INC.  
XX  
PI Choi GH, Bailey C, Hromockyj A, Kunsch CA;  
XX  
DR WPI: 2003-089120/08.  
DR P-PSDB; ABU13617.  
XX  
PT New EF040 polypeptides and polynucleotides from Enterococcus faecalis,  
PT useful for generating an immune response against E. faecalis and other  
PT Enterococcus species, and as vaccines against other bacterial genera.  
XX  
PS Example 1; Col 115-116; 146bp; English.  
XX  
CC The invention relates to polynucleotide fragments of a gene from  
CC Enterococcus faecalis, EF040, and the polypeptides encoded by them. The  
CC polypeptides are useful in detecting E. faecalis, as epitope tags, as  
CC molecular weight markers on SDS-PAGE gels or for molecular sieve gel  
CC filtration columns, in generating antibodies that specifically bind to  
CC the E. faecalis polypeptides, in generating an immune response against E.  
CC faecalis and other Enterococcus species and as vaccines against other  
CC bacterial genera. The polynucleotides are useful as probes for gene  
CC mapping and for identifying E. faecalis in biological samples. Sequences  
CC ABX1555-ABX61802 represent EF040 polynucleotides of the invention. Note:  
CC The sequence data for this patent can also be obtained from USPTO at  
CC segdata.uspto.gov/sequence.html  
XX  
SQ Sequence 1687 BP; 632 A; 284 C; 360 G; 411 T; 0 U; 0 Other;  
XX  
Query Match 89.7%; Score 1687; DB 8; Length 1687;  
Query Local Similarity 100.0%; Pred. No. 0;  
Matches 1687; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 96 AGAAGAAAATGGGAGAGCGACACGCTGTGATTTCAAAAAGAAATGACGATTTACC 155  
Db 1 AGAAGAAAATGGGAGAGCGACACGCTGTGATTTCAAAAAGAAATGACGATTTACC 60  
QY 156 AGATCCGCTTATTCAAATAGCGGGAAGAAATGAGCGAGTTGATTAATTAACAGGACT 215  
Db 61 AGATCCGCTTATTCAAATAGCGGGAAGAAATGAGCGAGTTGATTAATTAACAGGACT 120  
QY 216 GGCAGATGTGACGTTTATTAATTAATTAACGTCGACGAACTTTTACGACACGAGCGCG 275  
Db 121 GGCAGATGTGACGTTTATTAATTAATTAACGTCGACGAACTTTTACGACACGAGCGCG 180  
QY 276 AGGCGCAAGCGTTGATGACGCTTAAACAGCTGTCCAAAGTTTAATCTCTGGGAAACCTGT 335  
Db 181 AGGCGCAAGCGTTGATGACGCTTAAACAGCTGTCCAAAGTTTAATCTCTGGGAAACCTGT 240  
QY 336 TGCTCAAGGAACCAACGATGCAATGGAATGCAATGCTGTCAGTACCTTAAACCAAAA 395  
Db 241 TGCTCAAGGAACCAACGATGCAATGGAATGCAATGCTGTCAGTACCTTAAACCAAAA 300  
QY 396 TGGTAAAGATGACGATGATTAACATTAAGAAAGAACCAAGAGGCTGTGATGCTGTAC 455  
Db 301 TGGTAAAGATGACGATGATTAACATTAAGAAAGAACCAAGAGGCTGTGATGCTGTAC 360  
QY 456 GAATATGCTGTGCTGCTTCCAGTTTACGAATGATCAAGCAACAGATGCTTCTATA 515  
Db 361 GAATATGCTGTGCTGCTTCCAGTTTACGAATGATCAAGCAACAGATGCTTCTATA 420  
QY 516 ATATGGAACAGAAATTAAGCGGTTGTCATATTAATCTTAAATGAGTACCAATGA 575  
Db 421 ATATGGAACAGAAATTAAGCGGTTGTCATATTAATCTTAAATGAGTACCAATGA 480  
QY 576 TGGTATTAATGATGAAAAAAGTGAAGTGTGAAATGAAGATTAATGCGCGAGA 635  
Db 481 TGGTATTAATGATGAAAAAAGTGAAGTGTGAAATGAAGATTAATGCGCGAGA 540  
QY 636 ATTGTTATTTCTTAAAGCGAAGGCTCACAGGCACTGAATTAATTCAGAGAGTCAA 695

```

Db      541 ATTGTTATTTCTAAAAAGCGAAGCTCACGAGGACAGTAAATATATCATGAGAGTCAA 600
Qy      636 AGATGATATATATCATGAGCAACGATTAAGAACAGCAAAACGTTTATTAAGTGGAA 755
Db      601 AGATGATATATATCATGAGCAACGATTAAGAACAGCAAAACGTTTATTAAGTGGAA 660
Qy      756 AAGTTATGAAATTTGGGAAATGATTTTCAGAGAGGAAATGGAACGGGAGAAATTAAC 815
Db      661 AAGTTATGAAATTTGGGAAATGATTTTCAGAGAGGAAATGGAACGGGAGAAATTAAC 720
Qy      816 AGTTAAAAATCTTAGAGTTGGTTCGTATATTTAGAGAGATTAAGTCCAAATTAATGC 875
Db      721 AGTTAAAAATCTTAGAGTTGGTTCGTATATTTAGAGAGATTAAGTCCAAATTAATGC 780
Qy      876 AGAATTTAATTTGAAAAATCAACAAAAACACATTTTACATTTGAAGCAACATCAACAC 935
Db      781 AGAATTTAATTTGAAAAATCAACAAAAACACATTTTACATTTGAAGCAACATCAACAC 840
Qy      936 TGTGAAAAAAGATCAAAATGATACCTCTAAAGTTGATTAACAAACCAAGCTTAGA 995
Db      841 TGTGAAAAAAGATCAAAATGATACCTCTAAAGTTGATTAACAAACCAAGCTTAGA 900
Qy      996 TGTAAAGATGTGGCAATTTGGCAAAATTAATATCAATTTCTGTAAATATTCATTT 1055
Db      901 TGTAAAGATGTGGCAATTTGGCAAAATTAATATCAATTTCTGTAAATATTCATTT 960
Qy      1056 GGGGATTTGAGACAAAGAAAGGCGCTAATATACGTCAATTCATTTAGTTGATTA 1115
Db      961 GGGGATTTGAGACAAAGAAAGGCGCTAATATACGTCAATTCATTTAGTTGATTA 1020
Qy      1116 ACATGATGAGCGCTTAACCTTTGATTAACGTGACCTTGAGAGATAGCTTAAGCTTAA 1175
Db      1021 ACATGATGAGCGCTTAACCTTTGATTAACGTGACCTTGAGAGATAGCTTAAGCTTAA 1080
Qy      1176 TGATGGGATACAGTGAATGCTCTGAAATTAATCAAGTGAACAGCAATAGGCTT 1235
Db      1081 TGATGGGATACAGTGAATGCTCTGAAATTAATCAAGTGAACAGCAATAGGCTT 1140
Qy      1236 CACTGTCCGCGTTAATCCAGCGTATATTCCTACGCTTAACCAAGGCGGACACTTAAT 1295
Db      1141 CACTGTCCGCGTTAATCCAGCGTATATTCCTACGCTTAACCAAGGCGGACACTTAAT 1200
Qy      1296 CGTTACTTATGATATTTAAATGAAAGACAGATCCTACGAAAGGCTTAAATAGAGC 1355
Db      1201 CGTTACTTATGATATTTAAATGAAAGACAGATCCTACGAAAGGCTTAAATAGAGC 1260
Qy      1356 GAATGTGATTAACGCTCATACGACGACCAACCAACCACTGTTGAAGTTGTGACAG 1415
Db      1261 GAATGTGATTAACGCTCATACGACGACCAACCAACCACTGTTGAAGTTGTGACAG 1320
Qy      1416 TGGGAAACGTTTATTAAGTGTGATGAGGACGACCAAGCTTGGCGGAGC 1475
Db      1321 TGGGAAACGTTTATTAAGTGTGATGAGGACGACCAAGCTTGGCGGAGC 1380
Qy      1476 TTCCTTTGCTCCGCGTCAAAACAGGACGACGAAATTTTGAATATGATGAAC 1535
Db      1381 TTCCTTTGCTCCGCGTCAAAACAGGACGACGAAATTTTGAATATGATGAAC 1440
Qy      1536 AACGAAGACGACACTTGGGTGAAAAACAAGCTGAGCACTTCTTTTACAACAGCGC 1595
Db      1441 AACGAAGACGACACTTGGGTGAAAAACAAGCTGAGCACTTCTTTTACAACAGCGC 1500
Qy      1596 TGAATGATTAAGTATATCAACGAGGCTTAAATACGCTACCTATTTTGAAGAACTGT 1655
Db      1501 TGAATGATTAAGTATATCAACGAGGCTTAAATACGCTACCTATTTTGAAGAACTGT 1560
Qy      1656 AGCTCCGATGATATATGCTTTGTAAACAATCGAATGAAATGAGTCAATGACAATC 1715
Db      1561 AGCTCCGATGATATATGCTTTGTAAACAATCGAATGAAATGAGTCAATGACAATC 1620
Qy      1716 ATATGGCAACAGAAAACTAGATTTCACAGAAAAAGTACCAACCAACCAAGGATC 1775
Db      1621 ATATGGCAACAGAAAACTAGATTTCACAGAAAAAGTACCAACCAACCAAGGATC 1680

```

```

Qy      1776 CTTACTT 1782
Db      1681 CTTACTT 1687

RESULT 15
ADY38953
ID ADY38953 Standard; DNA; 1687 BP.
XX
AC ADY38953;
XX
DT 05-MAY-2005 (first entry)
XX
DE Novel Enterococcus faecalis gene sequence SeqID219.
XX
KW protein purification; DNA purification; antibacterial; vaccine;
KW enterococcus faecalis infection; gene; db.
XX
OS Enterococcus faecalis.
XX
PN US2005043528-A1.
XX
PD 24-FEB-2005.
XX
PF 06-AUG-2004; 2004US-00912362.
XX
PR 06-MAY-1997; 97US-0044031P.
XX
PR 16-MAY-1997; 97US-0046555P.
XX
PR 14-NOV-1997; 97US-0066009P.
XX
PR 04-MAY-1998; 98US-00071035.
XX
PR 29-JUL-2002; 2002US-00206576.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Choi GH, Bailey C, Hromockyj A, Kunsch CA;
XX
DR MPI: 2005-180840/19.
XX
DR P-PSDB; ADY38954.
XX
PT New isolated Enterococcus faecalis nucleic acid useful for diagnosing,
XX
PT preventing and/or attenuating infection by Enterococcus faecalis in
XX
PT animals, particularly in humans.
XX
PS Claim 1; SEQ ID NO 219; 28bp; English.
XX
CC This invention relates to novel genes and the proteins they encode
CC isolated from Enterococcus faecalis. The invention may be useful for the
CC development of compounds with an antibacterial activity or a vaccine. The
CC present invention is useful for the diagnosis, prevention and/or
CC attenuation of infection by Enterococcus faecalis in animals,
CC particularly in humans. The present sequence is that of a novel
CC Enterococcus faecalis gene of the invention. Note: The present sequence
CC does not form part of the printed specification but was obtained in
CC electronic form from the USPTO web site.
XX
SQ Sequence 1687 BP; 632 A; 284 C; 360 G; 411 T; 0 U; 0 Other;

Query Match      89.7%; Score 1687; DB 14; Length 1687;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1687; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      96 AGAAGAAAAATGGGAGAGCGCACAGCTGTGATTTCAAAAAAGAAATGACGGAATTAAC 155
Db      1 AGAAGAAAAATGGGAGAGCGCACAGCTGTGATTTCAAAAAAGAAATGACGGAATTAAC 60
Qy      156 AGATCCGCTTATTTCAAAATTAAGCGGAAAGAAATGAGCAAGTTGATTAATCAAGACT 215
Db      61 AGATCCGCTTATTTCAAAATTAAGCGGAAAGAAATGAGCAAGTTGATTAATCAAGACT 120
Qy      216 GGCAGATGTGACGTTTATTAATTAACGTGACGAACGAATTTTACGAGCAACGAGCGC 275
Db      121 GGCAGATGTGACGTTTATTAATTAATTAACGTGACGAACGAATTTTACGAGCAACGAGCGC 180

```

QY 276 AGGCGCAGCGTTGATGCAAGCTAACAAGCTGTCCAAAGTTTAACTCTCGGAAAACCTGT 335  
DB 181 AGGCGCAGCGTTGATGCAAGCTAACAAGCTGTCCAAAGTTTAACTCTCGGAAAACCTGT 240  
QY 336 TGCCTAAGGAACCAACGATGCAATGCGAATGTCACTGTTCACTTAACCTAAACAAACAA 395  
DB 241 TGCCTAAGGAACCAACGATGCAATGCGAATGTCACTGTTCACTTAACCTAAACAAACAA 300  
QY 396 TGTAAAGATGCACTGTATACCATTAAGAGAACCAAAAGGGGTGATGTTGCTGCTAC 455  
DB 301 TGTAAAGATGCACTGTATACCATTAAGAGAACCAAAAGGGGTGATGTTGCTGCTAC 360  
QY 456 GAATATGCTGTGGGCTTCCGATTTAGAAATGATCAAGCAAAAGATGTTCTATTA 515  
DB 361 GAATATGCTGTGGGCTTCCGATTTAGAAATGATCAAGCAAAAGATGTTCTATTA 420  
QY 516 ATATGCAACAGAAAGATTAGCGTGTTCATATTTATCTTAATAATGTTGATGCCAATGA 575  
DB 421 ATATGCAACAGAAAGATTAGCGTGTTCATATTTATCTTAATAATGTTGATGCCAATGA 480  
QY 576 TGTGATTTACATGTGAAAAAGTAGAACTGCTGAAATGAAAGTTTAATGCGCGAGA 635  
DB 481 TGTGATTTACATGTGAAAAAGTAGAACTGCTGAAATGAAAGTTTAATGCGCGAGA 540  
QY 636 ATTTGTTATTTCTAAGAGGCTCACAGGCAAGTAAATATATCAAGAGTCA 695  
DB 541 ATTTGTTATTTCTAAGAGGCTCACAGGCAAGTAAATATATCAAGAGTCA 600  
QY 696 AGATGATTTATATACATGCAACGATTAAGAAACAAAGCAAAACGTTTATTTACTGGGA 755  
DB 601 AGATGATTTATATACATGCAACGATTAAGAAACAAAGCAAAACGTTTATTTACTGGGA 660  
QY 756 AAGTATGAAATTTGCGAAAAATGATTTTCACAGACAGAGATGGAACGGGGAATTAAC 815  
DB 661 AAGTATGAAATTTGCGAAAAATGATTTTCACAGACAGAGATGGAACGGGGAATTAAC 720  
QY 816 AGTTAAAAATCTTGAGGTTGTTGTTATTTTGAAGAAAGTAAAGTCCAAATATATGC 875  
DB 721 AGTTAAAAATCTTGAGGTTGTTGTTATTTTGAAGAAAGTAAAGTCCAAATATATGC 780  
QY 876 AGAATTTAATTTGAAAAACAAACAAACACCATTTACATTTGAAGCAACATCAACACC 935  
DB 781 AGAATTTAATTTGAAAAACAAACAAACACCATTTACATTTGAAGCAACATCAACACC 840  
QY 936 TGTGAAAAAAACAGTCAAAAAATGATACCTTTAAGTTGATTAACAAACCAAGCTTAGA 995  
DB 841 TGTGAAAAAAACAGTCAAAAAATGATACCTTTAAGTTGATTAACAAACCAAGCTTAGA 900  
QY 996 TGTGAAAAAGTGTGGCAATGGGCAAAAAATTAATATCAATTTCTGTAAATATTTCCATT 1055  
DB 901 TGTGAAAAAGTGTGGCAATGGGCAAAAAATTAATATCAATTTCTGTAAATATTTCCATT 960  
QY 1056 GGGGATTTGCAACAAAGAAAGCGCTAATAAATACGTCAAAATTTCAATTTAGTTGATTA 1115  
DB 961 GGGGATTTGCAACAAAGAAAGCGCGCTAATAAATACGTCAAAATTTAGTTGATTA 1020  
QY 1116 ACATATGAGAGCTTAACTTTTGAATACCTGCTGAGAGTATGCTTAAGCGTTATA 1175  
DB 1021 ACATATGAGAGCTTAACTTTTGAATACCTGCTGAGAGTATGCTTAAGCGTTATA 1080  
QY 1176 TGAATGGGATACAGTGAATGCTCTGAAAAATTTATCAAGTGAACAAGCAAAATGGCTT 1235  
DB 1081 TGAATGGGATACAGTGAATGCTCTGAAAAATTTATCAAGTGAACAAGCAAAATGGCTT 1140  
QY 1236 CACTGTGCGCGTTAATCAAGCGTATATTCCTAAGCTAACAACAAGCGCGCAACTAAATT 1295  
DB 1141 CACTGTGCGCGTTAATCAAGCGTATATTCCTAAGCTAACAACAAGCGCGCAACTAAATT 1200  
QY 1296 CGTTTAACTTTATGCAATTTAAGAAACAGAGTCTTAAGAAAGGCTTTAAATGAGGC 1355  
DB 1201 CGTTTAACTTTATGCAATTTAAGAAACAGAGTCTTAAGAAAGGCTTTAAATGAGGC 1260

QY 1356 GAATGTTGATTAACGTTATACCGACGACCAACCAACCAACTGTTGAAGTTGTGACAGG 1415  
DB 1261 GAATGTTGATTAACGTTATACCGACGACCAACCAACCAACTGTTGAAGTTGTGACAGG 1320  
QY 1416 TGGGAAACGTTTCATTAAGTGTGATGCGGATGTGACAGGACCAAGCTTGGCGGAGC 1475  
DB 1321 TGGGAAACGTTTCATTAAGTGTGATGCGGATGTGACAGGACCAAGCTTGGCGGAGC 1380  
QY 1476 TTCCTTGTGCTGCGTGTATCAAAAACAGGCAACAGCAAAATTAATGAAATGATGAAC 1535  
DB 1381 TTCCTTGTGCTGCGTGTATCAAAAACAGGCAACAGCAAAATTAATGAAATGATGAAC 1440  
QY 1536 AACGAAACAGCAACTTGGGTGAAAAACAAAGCTGAACCACTACTTTTACAAACGCGC 1595  
DB 1441 AACGAAACAGCAACTTGGGTGAAAAACAAAGCTGAACCACTACTTTTACAAACGCGC 1500  
QY 1596 TGATGATTAAGTTGATATCAAGGCTTAAATACGGTACCTATTTTGAAGAAACCTGT 1655  
DB 1501 TGATGATTAAGTTGATATCAAGGCTTAAATACGGTACCTATTTTGAAGAAACCTGT 1560  
QY 1656 AGCTCTGATGATTAATGCTTGTAAACAATCGATTTGAATTTGTCAATGAACAATC 1715  
DB 1561 AGCTCTGATGATTAATGCTTGTAAACAATCGATTTGAATTTGTCAATGAACAATC 1620  
QY 1716 ATATGCAACAACGAAAACTTAGTTTCAACGAAAAAGTACCAACAAACAAAGGTAC 1775  
DB 1621 ATATGCAACAACGAAAACTTAGTTTCAACGAAAAAGTACCAACAAACAAAGGTAC 1680  
QY 1776 CTTTACT 1782  
DB 1681 CTTTACT 1687

Search completed: July 5, 2006, 18:22:45  
Job time : 832 secs

GenCore version 5.1.9  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 6, 2006, 00:18:42 ; Search time 1676 Seconds  
(without alignment)  
13790.598 Million cell updates/sec

Title: US-10-661-809A-12

Perfect score: 1881

Sequence: 1 atgaagcaatcaaaaaaagc.....gacgtagaagaagaatgct 1881

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 18892170 seqs, 6143817638 residues

37784340

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Published Applications NA Main:

1: /BMC\_Celerra\_SIDS3/ptodata/2/pubpna/US07\_PUBCOMB.seq:\*  
2: /BMC\_Celerra\_SIDS3/ptodata/2/pubpna/US08\_PUBCOMB.seq:\*  
3: /BMC\_Celerra\_SIDS3/ptodata/2/pubpna/US09\_PUBCOMB.seq:\*  
4: /BMC\_Celerra\_SIDS3/ptodata/2/pubpna/US09B\_PUBCOMB.seq:\*  
5: /BMC\_Celerra\_SIDS3/ptodata/2/pubpna/US09C\_PUBCOMB.seq:\*  
6: /BMC\_Celerra\_SIDS3/ptodata/2/pubpna/US10\_PUBCOMB.seq:\*  
7: /BMC\_Celerra\_SIDS3/ptodata/2/pubpna/US10B\_PUBCOMB.seq:\*  
8: /BMC\_Celerra\_SIDS3/ptodata/2/pubpna/US10C\_PUBCOMB.seq:\*  
9: /BMC\_Celerra\_SIDS3/ptodata/2/pubpna/US10D\_PUBCOMB.seq:\*  
10: /BMC\_Celerra\_SIDS3/ptodata/2/pubpna/US10E\_PUBCOMB.seq:\*  
11: /BMC\_Celerra\_SIDS3/ptodata/2/pubpna/US10F\_PUBCOMB.seq:\*  
12: /BMC\_Celerra\_SIDS3/ptodata/2/pubpna/US10G\_PUBCOMB.seq:\*  
13: /BMC\_Celerra\_SIDS3/ptodata/2/pubpna/US11A\_PUBCOMB.seq:\*  
14: /BMC\_Celerra\_SIDS3/ptodata/2/pubpna/US11B\_PUBCOMB.seq:\*  
15: /BMC\_Celerra\_SIDS3/ptodata/2/pubpna/US11C\_PUBCOMB.seq:\*  
16: /BMC\_Celerra\_SIDS3/ptodata/2/pubpna/US11D\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score  | Query Match | Length | DB ID | Description          |
|------------|--------|-------------|--------|-------|----------------------|
| 1          | 1881   | 100.0       | 1881   | 8     | US-10-661-809-12     |
| 2          | 1881   | 100.0       | 1983   | 3     | US-09-071-035-217    |
| 3          | 1881   | 100.0       | 1983   | 6     | US-10-206-576-217    |
| 4          | 1881   | 100.0       | 1983   | 9     | US-10-912-362-217    |
| 5          | 1881   | 100.0       | 15614  | 3     | US-09-070-927A-45    |
| 6          | 1887   | 89.7        | 1687   | 3     | US-09-071-035-219    |
| 7          | 1687   | 89.7        | 1687   | 6     | US-10-206-576-219    |
| 8          | 1687   | 89.7        | 1687   | 9     | US-10-912-362-219    |
| 9          | 1014.6 | 53.9        | 1878   | 8     | US-10-661-809-18     |
| 10         | 595    | 31.6        | 2199   | 8     | US-10-333-002-28     |
| 11         | 67     | 3.6         | 1062   | 9     | US-10-425-115-120013 |
| 12         | 65     | 3.5         | 627    | 8     | US-10-021-323-9336   |
| 13         | 64.6   | 3.4         | 6609   | 7     | US-11-020-509-1      |
| 14         | 64.6   | 3.4         | 1029   | 9     | US-10-425-115-123086 |
| 15         | 64.2   | 3.4         | 14006  | 7     | US-10-311-455-1931   |
| 17         | 63.2   | 3.4         | 1509   | 8     | US-10-333-002-15     |

|   |    |      |     |       |    |                      |                    |
|---|----|------|-----|-------|----|----------------------|--------------------|
| C | 18 | 63   | 3.3 | 822   | 9  | US-10-363-345A-11443 | Sequence 31443, A  |
| C | 19 | 63   | 3.3 | 822   | 9  | US-10-363-345A-11444 | Sequence 31444, A  |
| C | 20 | 63   | 3.3 | 822   | 9  | US-10-363-345A-18409 | Sequence 38409, A  |
| C | 21 | 63   | 3.3 | 822   | 9  | US-10-363-345A-38410 | Sequence 38410, A  |
| C | 22 | 63   | 3.3 | 822   | 10 | US-10-363-483A-31443 | Sequence 31443, A  |
| C | 23 | 63   | 3.3 | 822   | 10 | US-10-363-483A-31444 | Sequence 31444, A  |
| C | 24 | 63   | 3.3 | 822   | 10 | US-10-363-483A-38409 | Sequence 38409, A  |
| C | 25 | 63   | 3.3 | 822   | 10 | US-10-363-483A-38410 | Sequence 38410, A  |
| C | 26 | 62.6 | 3.3 | 739   | 9  | US-10-425-115-55020  | Sequence 55020, A  |
| C | 27 | 62.4 | 3.3 | 1214  | 8  | US-10-425-115-10203  | Sequence 10203, A  |
| C | 28 | 61.8 | 3.3 | 9539  | 6  | US-10-239-676-52     | Sequence 52, Appl  |
| C | 29 | 61.8 | 3.3 | 9539  | 7  | US-10-240-453-54     | Sequence 54, Appl  |
| C | 30 | 61.6 | 3.3 | 516   | 3  | US-09-960-352-5785   | Sequence 5785, Ap  |
| C | 31 | 61.2 | 3.3 | 1137  | 9  | US-10-425-115-40312  | Sequence 40312, A  |
| C | 32 | 61   | 3.2 | 1430  | 9  | US-10-425-115-13239  | Sequence 13239, A  |
| C | 33 | 60.6 | 3.2 | 1065  | 16 | US-11-189-817-1      | Sequence 1, Appl   |
| C | 34 | 60.4 | 3.2 | 1039  | 9  | US-10-425-115-58802  | Sequence 58802, A  |
| C | 35 | 60.4 | 3.2 | 1243  | 9  | US-10-425-115-172717 | Sequence 172717, A |
| C | 36 | 60   | 3.2 | 1121  | 9  | US-10-425-115-115706 | Sequence 115706, A |
| C | 37 | 60   | 3.2 | 1297  | 9  | US-10-425-115-67687  | Sequence 67687, A  |
| C | 38 | 60   | 3.2 | 15674 | 7  | US-10-311-455-136    | Sequence 336, Appl |
| C | 39 | 60   | 3.2 | 15674 | 7  | US-10-240-465-30     | Sequence 30, Appl  |
| C | 40 | 59.8 | 3.2 | 5750  | 7  | US-10-311-455-1981   | Sequence 1981, Ap  |
| C | 41 | 59.8 | 3.2 | 5750  | 8  | US-10-221-714A-433   | Sequence 433, Appl |
| C | 42 | 59.4 | 3.2 | 556   | 9  | US-10-357-930-40082  | Sequence 40082, A  |
| C | 43 | 59.4 | 3.2 | 556   | 9  | US-10-357-930-40182  | Sequence 40182, A  |
| C | 44 | 59.4 | 3.2 | 556   | 9  | US-10-357-930-42124  | Sequence 42124, A  |
| C | 45 | 59.4 | 3.2 | 556   | 9  | US-10-357-930-43620  | Sequence 43620, A  |

## ALIGNMENTS

RESULT 1  
US-10-661-809-12  
; Sequence 12, Application US/10661809  
; Publication No. US20040101919A1  
; GENERAL INFORMATION:  
; APPLICANT: HOOK Magnus  
; TITLE OF INVENTION: BIOINFORMATIC METHOD FOR IDENTIFYING SURFACE-ANCHORED PROTEINS FR  
; FILE REFERENCE: POSITIVE BACTERIA  
; CURRENT APPLICATION NUMBER: US/10/661,809  
; CURRENT FILING DATE: 2003-09-15  
; PRIOR APPLICATION NUMBER: 60/410303  
; PRIOR FILING DATE: 2002-09-13  
; NUMBER OF SEQ ID NOS: 24  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 12  
; LENGTH: 1881  
; TYPE: DNA  
; ORGANISM: Staphylococcus epidermidis  
US-10-661-809-12

|                       |              |              |                                            |              |
|-----------------------|--------------|--------------|--------------------------------------------|--------------|
| Query Match           | 100.0%       | Score 1881;  | DB 8;                                      | Length 1881; |
| Best Local Similarity | 100.0%       | Pred. No. 0; |                                            |              |
| Matches 1881;         | Conservative | 0;           | Mismatches                                 | 0;           |
|                       |              |              | Indels                                     | 0;           |
|                       |              |              | Gaps                                       | 0;           |
| QY                    | 1            | ATGAAGCAATTA | AAAAAGTTGGTACACCGTTAGTCTGTACTTAATTTGCCACTT | 60           |
| DB                    | 1            | ATGAAGCAATTA | AAAAAGTTGGTACACCGTTAGTCTGTACTTAATTTGCCACTT | 60           |
| QY                    | 61           | TTCAAGATGTA  | TGGAACAACATGCAATTCGATTTGAGAAAGGGGAGCGCACAG | 120          |
| DB                    | 61           | TTCAAGATGTA  | TGGAACAACATGCAATTCGATTTGAGAAAGGGGAGCGCACAG | 120          |
| QY                    | 121          | CTCGATTTCA   | CAAAAATGACGATTTTACCAATTCGCTTATCAAAATGACCGG | 180          |
| DB                    | 121          | CTCGATTTCA   | CAAAAATGACGATTTTACCAATTCGCTTATCAAAATGACCGG | 180          |
| QY                    | 181          | AAAGAAATGAG  | CGATTTGATTAATTCAGAGCTGGCAATGACGTTTACTTAT   | 240          |
| DB                    | 181          | AAAGAAATGAG  | CGATTTGATTAATTCAGAGCTGGCAATGACGTTTACTTAT   | 240          |



241 AACGTGAGAACGAATTTTACGAGCAACGAGGCGAAGCGTGAATGACGCTAA 300  
241 AACGTGAGAACGAATTTTACGAGCAACGAGGCGAAGCGTGAATGACGCTAA 300  
301 CAAGCTGTCACAAAGTTTAACTCTGGGAAACCTGTTGCTCAAGAAACCAACGATGCAAA 360  
301 CAAGCTGTCACAAAGTTTAACTCTGGGAAACCTGTTGCTCAAGAAACCAACGATGCAAA 360  
361 GGGAAATGTCAGTGTCTACGTTACCTTAAAAAATAATGTAAGATGCAATGTAATCCATT 420  
361 GGGAAATGTCAGTGTCTACGTTACCTTAAAAAATAATGTAAGATGCAATGTAATCCATT 420  
421 AAAAGAAACCAAAAGAGGAGTGAAGTGTGCTACGAATATGATGTCGCTCCCAAGT 480  
421 AAAAGAAACCAAAAGAGGAGTGAAGTGTGCTACGAATATGATGTCGCTCCCAAGT 480  
481 TACGAATATGATGTCAGTGTGCTACGAATATGATGTCGCTCCCAAGT 540  
481 TACGAATATGATGTCAGTGTGCTACGAATATGATGTCGCTCCCAAGT 540  
541 GTTCAATATTTATCTTAAAAATGTTGAGCAATGATGTTTACATGTAAGAAAGTA 600  
541 GTTCAATATTTATCTTAAAAATGTTGAGCAATGATGTTTACATGTAAGAAAGTA 600  
601 GGAATGCTGAAATTAAGAGATTTAAATGCGCGAAGATTTGTTATTTCTTAAAGCGAAGC 660  
601 GGAATGCTGAAATTAAGAGATTTAAATGCGCGAAGATTTGTTATTTCTTAAAGCGAAGC 660  
661 TCACGAGGACAGTAATATATATCCAAAGAGTCAAGATGATGATTAATCATGACAAAG 720  
661 TCACGAGGACAGTAATATATATCCAAAGAGTCAAGATGATGATTAATCATGACAAAG 720  
721 GATTAAGAAACAAAGCAAAACGCTTTATTTACTGGGAAAAGTTATGAAATTCGCAAAATGAT 780  
721 GATTAAGAAACAAAGCAAAACGCTTTATTTACTGGGAAAAGTTATGAAATTCGCAAAATGAT 780  
781 TTCACGAGGACAGTAATATATATCCAAAGAGTCAAGATGATGATTAATCATGACAAAG 840  
781 TTCACGAGGACAGTAATATATATCCAAAGAGTCAAGATGATGATTAATCATGACAAAG 840  
841 TATATTTTGAAGAAGTAATGAAGCTCCAATTAATGCAATTAATGAATCAACAAA 900  
841 TATATTTTGAAGAAGTAATGAAGCTCCAATTAATGCAATTAATGAATCAACAAA 900  
901 ACACCAATTTTACCAATGAGCAAAACATCAACACCTGTTGAAAACCACTCAAAATGAT 960  
901 ACACCAATTTTACCAATGAGCAAAACATCAACACCTGTTGAAAACCACTCAAAATGAT 960  
961 ACCCTTAAAGTGAATTAACCAACCAACCACTTGAATGATGAAGATGTCGCAATGCGCA 1020  
961 ACCCTTAAAGTGAATTAACCAACCAACCACTTGAATGATGAAGATGTCGCAATGCGCA 1020  
1021 AAAATTAATTAATCAATTTCTGTAATATATCCATTTGGGATGTCAGACAAAGAGCGAC 1080  
1021 AAAATTAATTAATCAATTTCTGTAATATATCCATTTGGGATGTCAGACAAAGAGCGAC 1080  
1081 GCTAATTAATTAATCAATTTCTGTAATATATCCATTTGGGATGTCAGACAAAGAGCGAC 1140  
1081 GCTAATTAATTAATCAATTTCTGTAATATATCCATTTGGGATGTCAGACAAAGAGCGAC 1140  
1141 AACGTGAGAACGAATTTTACGAGCAACGAGGCGAAGCGTGAATGACGCTAA 1200  
1141 AACGTGAGAACGAATTTTACGAGCAACGAGGCGAAGCGTGAATGACGCTAA 1200  
1201 GAAATTTATCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1260  
1201 GAAATTTATCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1260  
1261 ATTCTACGCTAACACGAGGCGACACTTAAATTCGTTTACTTATGCAATTTAAATGAA 1320  
1261 ATTCTACGCTAACACGAGGCGACACTTAAATTCGTTTACTTATGCAATTTAAATGAA 1320

1321 AAAGCAGATCTTACGAAAGCTTTAAATATGAGGCAATGTTGATTAACGCTCATACGAC 1380  
1321 AAAGCAGATCTTACGAAAGCTTTAAATATGAGGCAATGTTGATTAACGCTCATACGAC 1380  
1381 GACCAACACCAACCAACCTGTTGAAGTTGACAGGTGGGAAAAGCTTTCAATTAAGTCAT 1440  
1381 GACCAACACCAACCAACCTGTTGAAGTTGACAGGTGGGAAAAGCTTTCAATTAAGTCAT 1440  
1441 GCGCATGTCAGACGACCAAGCCCTTGGCGGAGCTTCTTGTGTCGCGATCAAAAC 1500  
1441 GCGCATGTCAGACGACCAAGCCCTTGGCGGAGCTTCTTGTGTCGCGATCAAAAC 1500  
1501 AGCGACAGCAAAATTTATTTGAAATCGATGAAAACAGAAAGCAGCAACTTGGTGAAA 1560  
1501 AGCGACAGCAAAATTTATTTGAAATCGATGAAAACAGAAAGCAGCAACTTGGTGAAA 1560  
1561 ACAAAGCTGAAGCAACTTCTTTTACCAACACGCTGATGATGATGATGATGATGATGATG 1620  
1561 ACAAAGCTGAAGCAACTTCTTTTACCAACACGCTGATGATGATGATGATGATGATGATG 1620  
1621 CTTAAATACGATCTTATTTTGAAGAACTGATGCTCTGATGATGATGATGATGATGATG 1680  
1621 CTTAAATACGATCTTATTTTGAAGAACTGATGCTCTGATGATGATGATGATGATGATG 1680  
1681 ACAATCGAATTTGATTTGCTCAATGACATCATATGACACAAAGAAACCTAGTT 1740  
1681 ACAATCGAATTTGATTTGCTCAATGACATCATATGACACAAAGAAACCTAGTT 1740  
1741 TCACGAGAAAGTAACCAACCAACCAAGGATCTTACCTTCAACAGGTGGCAAGAA 1800  
1741 TCACGAGAAAGTAACCAACCAACCAAGGATCTTACCTTCAACAGGTGGCAAGAA 1800  
1801 ATCTACGTTTCTTATGAGAGGCGGAGCTTCTGCTACTTATTTGAGAGGCTACTTGTCT 1860  
1801 ATCTACGTTTCTTATGAGAGGCGGAGCTTCTGCTACTTATTTGAGAGGCTACTTGTCT 1860  
1861 AGACGTAGAAAAAGAAATGCT 1881  
1861 AGACGTAGAAAAAGAAATGCT 1881

RESULT 2  
US-09-071-035-217  
Sequence 217, Application US/09071035  
Publication No. US20020045737A1  
GENERAL INFORMATION:  
APPLICANT: Gil H. Choi  
TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides  
NUMBER OF SEQUENCES: 496  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Human Genome Sciences, Inc.  
STREET: 9410 Key West Avenue  
CITY: Rockville  
STATE: Maryland  
COUNTRY: USA  
ZIP: 20850  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage  
COMPUTER: HP Vectra 486/33  
OPERATING SYSTEM: MSDOS version 6.2  
SOFTWARE: ASCII Text  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/071,035  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: A. Anders Brookes  
REGISTRATION NUMBER: 36,373  
REFERENCE/DOCKET NUMBER: PB369P2  
TELECOMMUNICATION INFORMATION:



TELEPHONE: (301) 309-8504  
TELEFAX: (301) 309-8512  
INFORMATION FOR SEQ ID NO: 217:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1983 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
US-09-071-035-217

Query Match 100.0%; Score 1881; DB 3; Length 1983;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1881; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 ATGAAGCATTTAAAAAGTTGGTACACCGTTAGTACCTTGTACTATTTTGGCAGTT 60
DB 100 ATGAAGCATTTAAAAAGTTGGTACACCGTTAGTACCTTGTACTATTTTGGCAGTT 159
QY 61 TTCACAGTGTATTAGGACAACTGCTGATTTGAGAGAAATGGGGAGCGCAG 120
DB 160 TTCACAGTGTATTAGGACAACTGCTGATTTGAGAGAAATGGGGAGCGCAG 219
QY 121 CTCGTGATTCACAAAAAGAAAATGACGATTTACAGATCCGCTTATTCAAAAATGCGGG 180
DB 220 CTCGTGATTCACAAAAAGAAAATGACGATTTACAGATCCGCTTATTCAAAAATGCGGG 279
QY 181 AAGAAATGAGCGAATTTGATTAATTTCAAGACTGGCAGATGTGACGTTTATTTAT 240
DB 280 AAGAAATGAGCGAATTTGATTAATTTCAAGACTGGCAGATGTGACGTTTATTTAT 339
QY 241 AACGTGACGAACGAATTTTACGAGCAAGGCGGCGAGGCGCAAGGCTTATGACAGTAA 300
DB 340 AACGTGACGAACGAATTTTACGAGCAAGGCGGCGAGGCGCAAGGCTTATGACAGTAA 399
QY 301 CAAGCTGTCCAAAGTTTAACTCTGGGAAACCTGTGCTCAAGAAACCAACCGATGCAAT 360
DB 400 CAAGCTGTCCAAAGTTTAACTCTGGGAAACCTGTGCTCAAGAAACCAACCGATGCAAT 459
QY 361 GGGATGTCACTGTTCACTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 420
DB 460 GGGATGTCACTGTTCACTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 519
QY 421 AAGAAAGAACCAAAAGAGGTGATGCTGCTAGCAATATGATGATGCTGCTGCTGCTGCT 480
DB 520 AAGAAAGAACCAAAAGAGGTGATGCTGCTAGCAATATGATGATGCTGCTGCTGCTGCT 579
QY 541 TACGAATGATCAACCAACAGATGTTCTTAAATATGAAACAGAAATTAAGCGTT 540
DB 580 TACGAATGATCAACCAACAGATGTTCTTAAATATGAAACAGAAATTAAGCGTT 639
QY 541 GTTCATATTATCTTAAATAATGTGTAGCAATGATGTTTACATGTGAAAAAGTA 600
DB 640 GTTCATATTATCTTAAATAATGTGTAGCAATGATGTTTACATGTGAAAAAGTA 699
QY 601 GGAATGCTGAAAATGAGAGATTAAATGCGCAGAAATTTGTAATTTCTTAAAGCGAAGC 660
DB 700 GGAATGCTGAAAATGAGAGATTAAATGCGCAGAAATTTGTAATTTCTTAAAGCGAAGC 759
QY 661 TCACACGACACAGTAAATATATTCAGAGAGTCAAGATGATTTATATCATGACAAAG 720
DB 760 TCACACGACACAGTAAATATATTCAGAGAGTCAAGATGATTTATATCATGACAAAG 819
QY 721 GATTAAGAAACAGCAAAACGCTTTATTTACTGAGAAAGTTATGAAATGGCGAAATGAT 780
DB 820 GATTAAGAAACAGCAAAACGCTTTATTTACTGAGAAAGTTATGAAATGGCGAAATGAT 879
QY 781 TTCACAGAGCAGAGATGAGACGAGAAATTAACAGTTAAATCTTGAGGTTGCTTCG 840
DB 880 TTCACAGAGCAGAGATGAGACGAGAAATTAACAGTTAAATCTTGAGGTTGCTTCG 939
QY 841 TATATTTTGAAGAGATTAAGCTCCAAATATGAGAAATTAATGAAATCAACAAA 900
DB 940 TATATTTTGAAGAGATTAAGCTCCAAATATGAGAAATTAATGAAATCAACAAA 999

```

```

QY 901 ACACGATTTACATTTGAAGCAACATCAACACCTGTTGAAAAAAGTCAAAAATGAT 960
DB 1000 ACACGATTTACATTTGAAGCAACATCAACACCTGTTGAAAAAAGTCAAAAATGAT 1059
QY 961 ACCCTTAAGTTGATTAACCAACACCAAGCTTGAATGTAAGATGTGCAATGGCGAA 1020
DB 1060 ACCCTTAAGTTGATTAACCAACACCAAGCTTGAATGTAAGATGTGCAATGGCGAA 1119
QY 1021 AAAATTAATCAAAATTTCTGTAAATATTCATTTGGGATTCAGACAAAGAGCGAC 1080
DB 1120 AAAATTAATCAAAATTTCTGTAAATATTCATTTGGGATTCAGACAAAGAGCGAC 1179
QY 1081 GCTAATTAATAGTCAAAATTTCAATTTAGTTGTAACATGATGACGCTTAACTTTGAT 1140
DB 1180 GCTAATTAATAGTCAAAATTTCAATTTAGTTGTAACATGATGACGCTTAACTTTGAT 1239
QY 1141 AACGTGACTTCGAGAGATATGCTTATGCTATATGATGAGGATATCAGTATTCCTCT 1200
DB 1240 AACGTGACTTCGAGAGATATGCTTATGCTTATATGATGAGGATATCAGTATTCCTCT 1299
QY 1201 GAAAATTAATCAAGTATGACCAACAGCAAAATGCTTCACTGTCGCGTTATTCAGCGTAT 1260
DB 1300 GAAAATTAATCAAGTATGACCAACAGCAAAATGCTTCACTGTCGCGTTATTCAGCGTAT 1359
QY 1261 ATTCTAGGCTTAAACCAAGCGGCGCACTAAATTCGTTACTTTATGATTTAAATGAA 1320
DB 1360 ATTCTAGGCTTAAACCAAGCGGCGCACTAAATTCGTTACTTTATGATTTAAATGAA 1419
QY 1321 AAGCAGATCTTACGAAAGGCTTTTAAATATGAGCGAAATGTTGATATCGGTATACCGAC 1380
DB 1420 AAGCAGATCTTACGAAAGGCTTTTAAATATGAGCGAAATGTTGATATCGGTATACCGAC 1479
QY 1381 GACCAACACCAACCACTGTTGAATGTGACAGGTGGAAAAGTTTCAATTAAGTCAT 1440
DB 1480 GACCAACACCAACCACTGTTGAATGTGACAGGTGGAAAAGTTTCAATTAAGTCAT 1539
QY 1441 GGGGATGTGACAGGACCAAGCCTTGGCGGAGGCTTCTTGTGCTCCGATCAAAAC 1500
DB 1540 GGGGATGTGACAGGACCAAGCCTTGGCGGAGGCTTCTTGTGCTCCGATCAAAAC 1599
QY 1501 AGCGACACAGCAAAATTTTGAATAATCGATGAAACAGAAAGCAACCTTGGGTGAA 1560
DB 1600 AGCGACACAGCAAAATTTTGAATAATCGATGAAACAGAAAGCAACCTTGGGTGAA 1659
QY 1561 ACAAAAGCTGAAGCACTTATTACAAACAGCTGATGATGATGATGATGATGATGATGAT 1620
DB 1660 ACAAAAGCTGAAGCACTTATTACAAACAGCTGATGATGATGATGATGATGATGATGAT 1719
QY 1621 CTTAAATAGGTAATCTTATTTAGAAAGAACTGATGCTCCGATGATGATGATGATGAT 1680
DB 1720 CTTAAATAGGTAATCTTATTTAGAAAGAACTGATGCTCCGATGATGATGATGATGAT 1779
QY 1681 ACAATTCGATTTGATTTGTGTCAATGAAACATATGGAACAAACAGAAACCTAGTT 1740
DB 1780 ACAATTCGATTTGATTTGTGTCAATGAAACATATGGAACAAACAGAAACCTAGTT 1839
QY 1741 TCACAGAAAAGTATCAACAAACAAAGGTAACCTTACCTTCAACAGGTGGCAAAGAA 1800
DB 1840 TCACAGAAAAGTATCAACAAACAAAGGTAACCTTACCTTCAACAGGTGGCAAAGAA 1899
QY 1801 ATCTAGGTTTCTTATGAGAAAGGCGCAGTCTTGTCACTTATTTGAGAGGTCTACTTGT 1860
DB 1900 ATCTAGGTTTCTTATGAGAAAGGCGCAGTCTTGTCACTTATTTGAGAGGTCTACTTGT 1959
QY 1861 AGACGTAGAAAAGAAATGCT 1881
DB 1960 AGACGTAGAAAAGAAATGCT 1980

```

RESULT 3  
US-10-206-576-217  
; Sequence 217, Application us/10206576

```
/ Publication No. US20030017495A1
/ GENERAL INFORMATION:
/ APPLICANT: Choi et al.
/ TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides
/ NUMBER OF SEQUENCES: 497
/ CORRESPONDENCE ADDRESS:
/ ADDRESS: Human Genome Sciences, Inc.
/ STREET: 9410 Key West Avenue
/ CITY: Rockville
/ STATE: Maryland
/ COUNTRY: USA
/ ZIP: 20850
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: CD-R
/ OPERATING SYSTEM: Windows 98
/ SOFTWARE: ASCII Text
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/10/206,576
/ FILING DATE: 29-Jul-2002
/ CLASSIFICATION: <Unknown>
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 09/071,035
/ FILING DATE: 1998-05-04
/ APPLICATION NUMBER: US 60/046,655
/ FILING DATE: 1997-05-16
/ APPLICATION NUMBER: US 60/044,031
/ FILING DATE: 1997-05-06
/ APPLICATION NUMBER: US 60/066,009
/ FILING DATE: 1997-11-14
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Hyman, Mark J.
/ REGISTRATION NUMBER: 46,789
/ REFERENCE/DOCKET NUMBER: PB369PDI1
/ INFORMATION FOR SEQ ID NO: 217:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 1983 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: double
/ TOPOLOGY: linear
/ SEQUENCE DESCRIPTION: SEQ ID NO: 217:
US-10-206-576-217

Query Match      100.0%; Score 1881; DB 6; Length 1983;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1881; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1  ATGAAGCAATTAATAAAGTTGGTACACCGTTAGTACCTTGTTACTAATTTTGCACCTT
DB      100  ATGAAGCAATTAATAAAGTTGGTACACCGTTAGTACCTTGTTACTAATTTTGCACCTT
QY      61  TTCAACAAGTGATTTTGGGACAACAACCTGATTTTGCAGAAATAATGGGAGAGCGCACG
DB      159  TTCAACAAGTGATTTTGGGACAACAACCTGATTTTGCAGAAATAATGGGAGAGCGCACG
QY      120  TTCAACAAGTGATTTTGGGACAACAACCTGATTTTGCAGAAATAATGGGAGAGCGCACG
DB      219  TTCAACAAGTGATTTTGGGACAACAACCTGATTTTGCAGAAATAATGGGAGAGCGCACG
QY      121  CTCGGATTTCAAAAAGAAATGACGAGTTTACCAAGATCCGCTTATTAATAATGCGGG
DB      279  CTCGGATTTCAAAAAGAAATGACGAGTTTACCAAGATCCGCTTATTAATAATGCGGG
QY      240  AAAAGAAATGAGCGAGTTTGATTAATATCAAGACCTGGCAGATGTGACGTTTATTTAT
DB      339  AAAAGAAATGAGCGAGTTTGATTAATATCAAGACCTGGCAGATGTGACGTTTATTTAT
QY      300  AACGTGACGAACGAATTTTACGACGACGAGCGGCGCAAGCGCTTATGACAGTTAA
DB      399  AACGTGACGAACGAATTTTACGACGACGAGCGGCGCAAGCGCTTATGACAGTTAA
QY      360  CAAGCTGCTCAAAAGTTTAACTCTGGGAAACCTGTGTCAGAGAAACCAACGATGCAAT
DB      459  CAAGCTGCTCAAAAGTTTAACTCTGGGAAACCTGTGTCAGAGAAACCAACGATGCAAT
QY      420  GGGAAATGTCATGTTCAAGTTACTAATAAATAAATAATGTAAGATGCAATGATACAT
DB      420  GGGAAATGTCATGTTCAAGTTACTAATAAATAAATAATGTAAGATGCAATGATACAT
```

```
DB      519  GGGAAATGTCATGTTCAAGTTACTAATAAATAAATAATGTAAGATGCAATGATACAT
QY      480  AAAAGAAATGAGCGGTGCTAGTTCCTGTAGCAATTAAGTGTCGCTCCAGTT
DB      579  AAAAGAAATGAGCGGTGCTAGTTCCTGTAGCAATTAAGTGTCGCTCCAGTT
QY      540  TACGAAATGATCAACCAACGATGTTCTCTAATAATGGAACGAAGATTTACGGTT
DB      639  TACGAAATGATCAACCAACGATGTTCTCTAATAATGGAACGAAGATTTACGGTT
QY      600  GTTCATATTTATCTTAAATAATGTTAGCGCAATGATGTTTACATGTAAGAAAGTA
DB      699  GTTCATATTTATCTTAAATAATGTTAGCGCAATGATGTTTACATGTAAGAAAGTA
QY      660  GGAACCTGCTGAAAATGAAAGATTAAATGGCGGAGATTGTTATTTCTTAAAGCAAGGC
DB      759  GGAACCTGCTGAAAATGAAAGATTAAATGGCGGAGATTGTTATTTCTTAAAGCAAGGC
QY      720  TCACGAGCACAAGTAAATATATCCAGAGAGTCAAGATGATTTATATACATGACCAAG
DB      819  TCACGAGCACAAGTAAATATATCCAGAGAGTCAAGATGATTTATATACATGACCAAG
QY      780  GATTAAGAAACAAGCAAAACGCTTTATATCTGGGAAAAGTTATGAAATGGCGAAATGAT
DB      879  GATTAAGAAACAAGCAAAACGCTTTATATCTGGGAAAAGTTATGAAATGGCGAAATGAT
QY      840  TTCAAGAAAGCAAGAAATGGAACCGGGAATTTAAGTTAAATAATCTGAGTTGCTCG
DB      939  TTCAAGAAAGCAAGAAATGGAACCGGGAATTTAAGTTAAATAATCTGAGTTGCTCG
QY      900  TATATTTAGAAAGATTAAGCTTCAATATATGCAATTAATTAATGAAATCAACAAA
DB      999  TATATTTAGAAAGATTAAGCTTCAATATATGCAATTAATTAATGAAATCAACAAA
QY      960  ACACCATTTTACATTTGAAGCAAAACATCAACCTGTTGAAAAACGTCAAAATGAT
DB      1059  ACACCATTTTACATTTGAAGCAAAACATCAACCTGTTGAAAAACGTCAAAATGAT
QY      1020  ACCCTTAAGTTGATTAACCAACACCAAGCTTAGATGTTAAAGATGGGCAATGGCGGA
DB      1119  ACCCTTAAGTTGATTAACCAACACCAAGCTTAGATGTTAAAGATGGGCAATGGCGGA
QY      1080  AAAATTAATATCAATTTTCTGTAATATATTCATTTGGGATTTGCAGACAAAGAGCGAC
DB      1179  AAAATTAATATCAATTTTCTGTAATATATTCATTTGGGATTTGCAGACAAAGAGCGAC
QY      1140  GCTAATTAATACGTCAAATTCATTTAGTTGATTAACATGATGACGCTTAACTTTGAT
DB      1239  GCTAATTAATACGTCAAATTCATTTAGTTGATTAACATGATGACGCTTAACTTTGAT
QY      1200  AACGTGACCTTCGAGAGATGCTTATGCTTATGATGAGGAGTACAGTATGCTCCT
DB      1299  AACGTGACCTTCGAGAGATGCTTATGCTTATGATGAGGAGTACAGTATGCTCCT
QY      1260  GAAATTTATCAAGTACGTAACAACCAATATGCTTCACTGTCGCCGTTAATCAACGAT
DB      1359  GAAATTTATCAAGTACGTAACAACCAATATGCTTCACTGTCGCCGTTAATCAACGAT
QY      1320  ATTCTGACGCTTAACCAACGCGGCACTTAAATTCGTTTACTTTATGATTAATGAA
DB      1419  ATTCTGACGCTTAACCAACGCGGCACTTAAATTCGTTTACTTTATGATTAATGAA
QY      1380  AAAAGAAATCCACGAAGGCTTTAAATAATGAGGGAATGTTGATTAAGGTCATCCGAC
DB      1479  AAAAGAAATCCACGAAGGCTTTAAATAATGAGGGAATGTTGATTAAGGTCATCCGAC
QY      1440  GACCAAAACACCAACCTGTTGAAGTTGACAGGTGGGAAAACGTTTCAATTAAGATGAT
DB      1539  GACCAAAACACCAACCTGTTGAAGTTGACAGGTGGGAAAACGTTTCAATTAAGATGAT
QY      1500  GGGGATGTCAGCGACCAACAGCTTGGCGGGAAGCTTCTTGTGTCGCTGATCAAAAC
DB      1599  GGGGATGTCAGCGACCAACAGCTTGGCGGGAAGCTTCTTGTGTCGCTGATCAAAAC
```



Db 1060 ACCTTAAGTTGATTAACCAACCAAGCTTATGATGTAAGAGTGGCAATTTGGCGAA 1119  
Qy 1021 AAAATTAATCAAAATTTCTGTAATATTCATTTGGGGATTCAGACAAAGAGCGAC 1080  
Db 1120 AAAATTAATCAAAATTTCTGTAATATTCATTTGGGGATTCAGACAAAGAGCGAC 1179  
Qy 1081 GCTAATAAATAGTCAAAATTCATTTAGTGAATGAATGATGAGCGCTTAATCTTTGAT 1140  
Db 1180 GCTAATAAATAGTCAAAATTCATTTAGTGAATGAATGATGAGCGCTTAATCTTTGAT 1239  
Qy 1141 AACGTACTTCTGAGAGATATCTTATGCGTTATATGATGAGGATATCAAGTATGCTCCT 1200  
Db 1240 AACGTACTTCTGAGAGATATCTTATGCGTTATATGATGAGGATATCAAGTATGCTCCT 1299  
Qy 1201 GAAATTTATCAAGTATGATCAAGCAAAATGCGCTTCACTGTCGCGTTAATCAGCGTAT 1260  
Db 1300 GAAATTTATCAAGTATGATCAAGCAAAATGCGCTTCACTGTCGCGTTAATCAGCGTAT 1359  
Qy 1261 ATTCTAGCGCTAACCAGGCGGCGACACTAAATTCGTTATCTTTATGATTAATGA 1320  
Db 1360 ATTCTAGCGCTAACCAGGCGGCGACACTAAATTCGTTATCTTTATGATTAATGA 1419  
Qy 1321 AAAGCAGATCTTACGAAAGGCTTTAAATATGAGCGCATATGATTAACCGTATACCGAC 1380  
Db 1420 AAAGCAGATCTTACGAAAGGCTTTAAATATGAGCGCATATGATTAACCGTATACCGAC 1479  
Qy 1381 GACCAAAACCAACCAACTGTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1440  
Db 1480 GACCAAAACCAACCAACTGTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1539  
Qy 1441 GGGGATGATGACGCGGACCAAGCTTTGGGGGAGCTTCTTTGCGTCCGATGATCAAAAC 1500  
Db 1540 GGGGATGATGACGCGGACCAAGCTTTGGGGGAGCTTCTTTGCGTCCGATGATCAAAAC 1599  
Qy 1501 AGCGACACGCAAAATTTATGAAATCGATGAAACAAACGAAAGCAAGCACTTGGGTGAA 1560  
Db 1600 AGCGACACGCAAAATTTATGAAATCGATGAAACAAACGAAAGCAAGCACTTGGGTGAA 1659  
Qy 1561 ACAAAGCTGAAAGCACTCTTTTACAAACAAGCGCTGATGATGATGATGATGATGATGATGAT 1620  
Db 1660 ACAAAGCTGAAAGCACTCTTTTACAAACAAGCGCTGATGATGATGATGATGATGATGATGAT 1719  
Qy 1621 CTTAATATAGTACCTATTTATTAAGAAAGCTGATGCTCCGATGATGATGATGATGATGATGAT 1680  
Db 1720 CTTAATATAGTACCTATTTATTAAGAAAGCTGATGCTCCGATGATGATGATGATGATGATGAT 1779  
Qy 1681 ACAAATCGGATTTGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1740  
Db 1780 ACAAATCGGATTTGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1839  
Qy 1741 TCACCAAGAAAGTACCAACCAAGCAAGGATCTTACCTTCAACAGGTGGCAAGAA 1800  
Db 1840 TCACCAAGAAAGTACCAACCAAGCAAGGATCTTACCTTCAACAGGTGGCAAGAA 1899  
Qy 1801 ATCTAGCTTACTTGAAGAGTGGCGGATCTTGTCTACTATTTGACAGAGTCTACTTTGCT 1860  
Db 1900 ATCTAGCTTACTTGAAGAGTGGCGGATCTTGTCTACTATTTGACAGAGTCTACTTTGCT 1959  
Qy 1861 AGACGTAGAAAGAAATGCT 1881  
Db 1960 AGACGTAGAAAGAAATGCT 1980

RESULT 5  
US-09-070-927A-45  
; Sequence 45, Application US/09070927A  
; Patent No. US20020120116A1  
; GENERAL INFORMATION:  
; APPLICANT: Charles A. Kunsch  
; Patrick J. Dillon  
; TITLE OF INVENTION: Enterococcus faecialis Polynucleotides and Polypeptides

NUMBER OF SEQUENCES: 982  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Human Genome Sciences, Inc.  
STREET: 9410 Key West Avenue  
CITY: Rockville  
STATE: Maryland  
COUNTRY: USA  
ZIP: 20850  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.44Mb storage  
COMPUTER: HP Vectra 486/33  
OPERATING SYSTEM: MSDOS version 6.2  
SOFTWARE: ASCII Text  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/070,927A  
FILING DATE: 04-May-2000  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/046,655  
FILING DATE: 1997-05-16  
APPLICATION NUMBER: 60/044,031  
FILING DATE: 1997-05-06  
APPLICATION NUMBER: 60/066,009  
FILING DATE: 1997-11-14  
ATTORNEY/AGENT INFORMATION:  
NAME: Kenley K. Hoover  
REGISTRATION NUMBER: 40,302  
REFERENCE/DOCKET NUMBER: PB369  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (301) 309-8504  
TELEFAX: (301) 309-8512  
INFORMATION FOR SEQ ID NO: 45:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 15614 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
SEQUENCE DESCRIPTION: SEQ ID NO: 45:  
US-09-070-927A-45

Query Match 100.0%; Score 1881; DB 3; Length 15614;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1881; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGAAGCAATTAAGAAAGTTGGTACACCGTTAGTACCTTGTACTAATTTTGGCACTT 60  
Db 12164 ATGAAGCAATTAAGAAAGTTGGTACACCGTTAGTACCTTGTACTAATTTTGGCACTT 12223  
Qy 61 TTCACAACTGTATTAGGACAAACAATGCAATTTGCAAGAAATGGGAGAGCGCACAG 120  
Db 12224 TTCACAACTGTATTAGGACAAACAATGCAATTTGCAAGAAATGGGAGAGCGCACAG 12283  
Qy 121 CTCGTGATTCACAAAAGAAATGACGATTTTACAGATCCGCTTAATCAAAATATGCGGG 180  
Db 12284 CTCGTGATTCACAAAAGAAATGACGATTTTACAGATCCGCTTAATCAAAATATGCGGG 12343  
Qy 181 AAAAGAAATGAGGAGTTGATTAATATCAAGGACTGCGAGATGATGACGTTTATAGTAT 240  
Db 12344 AAAAGAAATGAGGAGTTGATTAATATCAAGGACTGCGAGATGATGACGTTTATAGTAT 12403  
Qy 241 AACGTAGCAAGCAATTTTACGAGCAACGAGCGGAGCGCAAGCGTTGATGACCTTAA 300  
Db 12404 AACGTAGCAAGCAATTTTACGAGCAACGAGCGGAGCGCAAGCGTTGATGACCTTAA 12463  
Qy 301 CAAGCTGTCCAAAGTTTAATCTCTGGGAAACCTGTTGCTCAAGGAAACCAAGGATGCAAT 360  
Db 12464 CAAGCTGTCCAAAGTTTAATCTCTGGGAAACCTGTTGCTCAAGGAAACCAAGGATGCAAT 12523  
Qy 361 GGAATGTCACTGTTCACTTACCTTAAATAAATAATGATTAAGATGATGATGATGATGATGAT 420  
Db 12524 GGAATGTCACTGTTCACTTACCTTAAATAAATAATGATTAAGATGATGATGATGATGATGAT 12583  
Qy 421 AAAGAAAGCAAAAGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480



QY 96 AGAAGAAAATGGGGAGAGCGCAGCTGCTGATTCACAAAAGAAAATGACGATTTAC 155  
DB 1 AGAAGAAAATGGGGAGAGCGCAGCTGCTGATTCACAAAAGAAAATGACGATTTAC 60  
QY 156 AGATCCGCTTATTCAAAATAGCGGGAGAAAATGAGCGAGTTGATTAATCAAGACT 215  
DB 61 AGATCCGCTTATTCAAAATAGCGGGAGAAAATGAGCGAGTTGATTAATCAAGACT 120  
QY 216 GGCACATGAGAGCTTATGATTTATTAACGTGACGAACGAATTTTACGACACGAGCGC 275  
DB 121 GGCACATGAGAGCTTATGATTTATTAACGTGACGAACGAATTTTACGACACGAGCGC 180  
QY 276 AGGCGCAAGCGCTGATGACGCTAAACAGCTGTCCAAAGTTTAACTCTGGGAAACTGT 335  
DB 181 AGGCGCAAGCGCTGATGACGCTAAACAGCTGTCCAAAGTTTAACTCTGGGAAACTGT 240  
QY 336 TGCTCAAGAACCCACCGATGCAAAATGGAAATGTCACTGTTCACTTAACTTAAACAAA 395  
DB 241 TGCTCAAGAACCCACCGATGCAAAATGGAAATGTCACTGTTCACTTAACTTAAACAAA 300  
QY 396 TGCTCAAGAACCCACCGATGCAAAATGGAAATGTCACTGTTCACTTAACTTAAACAAA 455  
DB 301 TGCTCAAGAACCCACCGATGCAAAATGGAAATGTCACTGTTCACTTAACTTAAACAAA 360  
QY 456 GAATATGCTGTGGCGCTCCAGTTTACGAATGATCAAGCAACAGATGTTCTCTATA 515  
DB 361 GAATATGCTGTGGCGCTCCAGTTTACGAATGATCAAGCAACAGATGTTCTCTATA 420  
QY 516 ATATGGAACAGAAAGATTAGCGGTTGCTATATTTATCTTAAATGTGTGACCAATGA 575  
DB 421 ATATGGAACAGAAAGATTAGCGGTTGCTATATTTATCTTAAATGTGTGACCAATGA 480  
QY 576 TGCTGATTTACATGGAAGAAAGTGAAGCTGCTGAAGTGAAGTTTAAATGGCGCAGA 635  
DB 481 TGCTGATTTACATGGAAGAAAGTGAAGCTGCTGAAGTGAAGTTTAAATGGCGCAGA 540  
QY 636 ATTTGTTATTTCTTAAAGGGAAGGCTCAACAGCAACAGTAAATATATCCAGAGACTCA 695  
DB 541 ATTTGTTATTTCTTAAAGGGAAGGCTCAACAGCAACAGTAAATATATCCAGAGACTCA 600  
QY 696 AGATGATTTATATCAATGACACGATTAAGAACAGCAAAACGTTTATTTACTGGGAA 755  
DB 601 AGATGATTTATATCAATGACACGATTAAGAACAGCAAAACGTTTATTTACTGGGAA 660  
QY 756 AAGTATGAAATTTGGGAAAATGATTTTCCAGAGACGAGAAATGAAACGGAGATTTAC 815  
DB 661 AAGTATGAAATTTGGGAAAATGATTTTCCAGAGACGAGAAATGAAACGGAGATTTAC 720  
QY 816 AGTTAAAAATCTTGAAGTTGTTGCTATATTTTGAAGAGTAAAGCTCCAAATATATGC 875  
DB 721 AGTTAAAAATCTTGAAGTTGTTGCTATATTTTGAAGAGTAAAGCTCCAAATATATGC 780  
QY 876 AGAATTTAATTTGAATTCAAAACAAACACCATTTTACATTTGAAGCAACATCAAAACAC 935  
DB 781 AGAATTTAATTTGAATTCAAAACAAACACCATTTTACAAATTTGAAGCAACATCAAAACAC 840  
QY 936 TGTGAAAAAAACAGTCAAAAATGATACCTCTTAAAGTTGATTAACCAACCAAGCTTGA 995  
DB 841 TGTGAAAAAAACAGTCAAAAATGATACCTCTTAAAGTTGATTAACCAACCAAGCTTGA 900  
QY 996 TGTGAAAAAGTGTGGCAATTTGGGAAAATTAATATCAAAATTTCTGTAATATTTCCATT 1055  
DB 901 TGTGAAAAAGTGTGGCAATTTGGGAAAATTAATATCAAAATTTCTGTAATATTTCCATT 960  
QY 1056 GGGGATTTGACACAAAGAAAGGCGCTAATTAATTCGCAATTTCAATTTAGTTGATTA 1115  
DB 961 GGGGATTTGACACAAAGAAAGGCGCTAATTAATTCGCAATTTCAATTTAGTTGATTA 1020  
QY 1116 ACATGATGACGCTTAACTTTTGAATTAAGTGAAGTTGATTAAGTTGATTAAGTTGATTA 1175  
DB 1021 ACATGATGACGCTTAACTTTTGAATTAAGTGAAGTTGATTAAGTTGATTAAGTTGATTA 1080  
QY 1176 TGATGGGATTAACGATGATTTCTCTGAAAAATTAATCAAGTGAACAAAGCAATGGCTT 1235

DB 1081 TGATGGGATTAACGATGATTTGCTCTGAAAAATTAATCAAGTGAACAAAGCAATGGCTT 1140  
QY 1236 CACTGTCCCGCTTAAATCCAGGATATATTCCTACGTAACACAGCGCGGCACTTAAATTT 1295  
DB 1141 CACTGTCCCGCTTAAATCCAGGATATATTCCTACGTAACACAGCGCGGCACTTAAATTT 1200  
QY 1296 CGTTTACTTTTATGCAATTTAAATGAAGAAAGAGATCTCTACGAAGAGCTTTAAATATGAGGC 1355  
DB 1201 CGTTTACTTTTATGCAATTTAAATGAAGAAAGAGATCTCTACGAAGAGCTTTAAATATGAGGC 1260  
QY 1356 GAATGTTGATTAACGCTCATACCGACGACCAACACCAACTGTTGAAGTTGACAGG 1415  
DB 1261 GAATGTTGATTAACGCTCATACCGACGACCAACACCAACTGTTGAAGTTGACAGG 1320  
QY 1416 TGGGAAGCTTTCAATTAAGTCGATGAGGAGTGAACAGCAACAAAGCTTGGCGGAGC 1475  
DB 1321 TGGGAAGCTTTCAATTAAGTCGATGAGGAGTGAACAGCAACAAAGCTTGGCGGAGC 1380  
QY 1476 TTCTTTGCTGCTCGATCAAAAACAGCGACACAGCAATTAATTTGAAATGATGAAC 1535  
DB 1381 TTCTTTGCTGCTCGATCAAAAACAGCGACACAGCAATTAATTTGAAATGATGAAC 1440  
QY 1536 AACGAACGACCACTTGGGTGAAGAAACAAAGCTGAAGCACTTTTACACACGCGC 1595  
DB 1441 AACGAACGACCACTTGGGTGAAGAAACAAAGCTGAAGCACTTTTACACACGCGC 1500  
QY 1596 TGATGATTTAGTTGATTAACAGGGCTTAAATAGCGTCACTTATTTTGAAGAAACTGT 1655  
DB 1501 TGATGATTTAGTTGATTAACAGGGCTTAAATAGCGTCACTTATTTTGAAGAAACTGT 1560  
QY 1656 AGCTCTGATGATTAATGCTTGTGTTAAACAAATCGAATGAAATTTGTGCAATGAACATC 1715  
DB 1561 AGCTCTGATGATTAATGCTTGTGTTAAACAAATCGAATGAAATTTGTGCAATGAACATC 1620  
QY 1716 ATATGCGCAACAGAAAACCTTATGTTTACACGAAAAAGTACCAACAAACAGAGTAC 1775  
DB 1621 ATATGCGCAACAGAAAACCTTATGTTTACACGAAAAAGTACCAACAAACAGAGTAC 1680  
QY 1776 CTTACT 1782  
DB 1681 CTTACT 1687

RESULT 7  
US-10-206-576-219  
; Sequence 219, Application US/10206576  
; Publication No. US20030017495A1  
; GENERAL INFORMATION:  
; APPLICANT: Choi et al.  
; TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides  
; NUMBER OF SEQUENCES: 497  
; CORRESPONDENCE ADDRESS:  
; ADDRESSER: Human Genome Sciences, Inc.  
; STREET: 9410 Key West Avenue  
; CITY: Rockville  
; STATE: Maryland  
; COUNTRY: USA  
; ZIP: 20850  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: CD-R  
; COMPUTER: Dell Latitude  
; OPERATING SYSTEM: Windows 98  
; SOFTWARE: ASCII Text  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/10/206,576  
; FILING DATE: 29-Jul-2002  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 09/071,035  
; FILING DATE: 1998-05-04  
; APPLICATION NUMBER: US 60/046,655  
; FILING DATE: 1997-05-16

APPLICATION NUMBER: US 60/044,031  
FILING DATE: 1997-05-06  
APPLICATION NUMBER: US 60/066,009  
FILING DATE: 1997-11-14  
ATTORNEY/AGENT INFORMATION:  
NAME: Hyman, Mark J.  
REGISTRATION NUMBER: 46,789  
REFERENCE/DOCKET NUMBER: PB369P1D1  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1687 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
SEQUENCE DESCRIPTION: SEQ ID NO: 219:  
US-10-206-576-219

Query Match 89.7%; Score 1687; DB 6; Length 1687;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1687; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 96 AGAAGAAAATGGGAGAGCGCACAGCTCGTGAATTCACAAAAGAAAATGACGATTTACC 155  
DB 1 AGAAGAAAATGGGAGAGCGCACAGCTCGTGAATTCACAAAAGAAAATGACGATTTACC 60  
QY 156 AGATCGCTTATTCAAAATAGGGGAGAAAAGAAATAGCGAGTTGATTAATATCAAGGACT 215  
DB 61 AGATCGCTTATTCAAAATAGCGGAGAAAAGAAATAGCGAGTTGATTAATATCAAGGACT 120  
QY 216 GGCAGATGACGCTTAGATTATTAACGTGACGAACGAATTTTACGACCAACGACGCG 275  
DB 121 GGCAGATGACGCTTAGATTATTAACGTGACGAACGAATTTTACGACCAACGACGCG 180  
QY 276 AGGCGCAAGCGTTGATGACGTAAACAGCTGTCCAAAGTTTAACTCTGGGAAACTGT 335  
DB 181 AGGCGCAAGCGTTGATGACGTAAACAGCTGTCCAAAGTTTAACTCTGGGAAACTGT 240  
QY 336 TGCTCAAGGAACCAACGATGCAAAATGGGAAATGTCAGTCTGTTCACTTAACCTAAACAA 395  
DB 241 TGCTCAAGGAACCAACGATGCAAAATGGGAAATGTCAGTCTGTTCACTTAACCTAAACAA 300  
QY 396 TGGTAAAGATGACGTATACCATTAAGAAACCAAAAAGAGGATGATGCTGCTAC 455  
DB 301 TGGTAAAGATGACGTATACCATTAAGAAACCAAAAAGAGGATGATGCTGCTAC 360  
QY 456 GAATATGATGATGCGGTTCCAGTTTACGAAATGATCAACCAACGATGTTCTATTA 515  
DB 361 GAATATGATGATGCGGTTCCAGTTTACGAAATGATCAACCAACGATGTTCTATTA 420  
QY 516 ATATGGAACAGAAATTAAGCGGTTGTCATATTTATCCTAAATAATGTGTGACCAATGA 575  
DB 421 ATATGGAACAGAAATTAAGCGGTTGTCATATTTATCCTAAATAATGTGTGACCAATGA 480  
QY 576 TGGTATTTTACATGTAAGAAAGTAGAACTGCTGAATTAAGATTAATAGCGCAGA 635  
DB 481 TGGTATTTTACATGTAAGAAAGTAGAACTGCTGAATTAAGATTAATAGCGCAGA 540  
QY 636 ATTTGTTATTTCTAAAAGGAGGCTCACAGGACACAGTAAATATATCAAGAGTCA 695  
DB 541 ATTTGTTATTTCTAAAAGGAGGCTCACAGGACACAGTAAATATATCAAGAGTCA 600  
QY 696 AGATGATTTATTAATGACGATTAAGAAACAAAGAAAACGTTTATTAATCTGGGAA 755  
DB 601 AGATGATTTATTAATGACGATTAAGAAACAAAGAAAACGTTTATTAATCTGGGAA 660  
QY 756 AAGTTATGAATTTGGCGAAATGATTTCAAGAGGAGAAATGAAACGGAGAAATTAAC 815  
DB 661 AAGTTATGAATTTGGCGAAATGATTTCAAGAGGAGAAATGAAACGGAGAAATTAAC 720  
QY 816 AGTTAAATCTTGAAGTTGTTGTTATTTTAAAGAAAGTAAAGCTTCAATTAATGC 875  
DB 721 AGTTAAATCTTGAAGTTGTTGTTATTTTAAAGAAAGTAAAGCTTCAATTAATGC 780

QY 876 AGAATTAATTTGAATCAACAAACAAACCAATTAATTAATGAACAAACAAACACC 935  
DB 781 AGAATTAATTTGAATCAACAAACAAACCAATTAATTAATGAACAAACAAACACC 840  
QY 936 TGTGAAAAACAGTCAAAAATGATACCTCTAAAGTTGATTAACCAACCAAGCTTAA 995  
DB 841 TGTGAAAAACAGTCAAAAATGATACCTCTAAAGTTGATTAACCAACCAAGCTTAA 900  
QY 996 TGTGAAAGATGAGCAATTTGGGAAAAAATTAATATCAAAATTTCTGTAATATTCATT 1055  
DB 901 TGTGAAAGATGAGCAATTTGGGAAAAAATTAATATCAAAATTTCTGTAATATTCATT 960  
QY 1056 GGGGATTTGACAAAGAAAGGCGAGCTTAATTAATACGTAATTAATTAATGATTA 1115  
DB 961 GGGGATTTGACAAAGAAAGGCGAGCTTAATTAATACGTAATTAATTAATGATTA 1020  
QY 1116 ACATGATGACGCTTAATCTTTGATTAACGTGACTTCTGGAGATATGCTTAATGCTTAA 1175  
DB 1021 ACATGATGACGCTTAATCTTTGATTAACGTGACTTCTGGAGATATGCTTAATGCTTAA 1080  
QY 1176 TGAATGGGATACAGATTTGCTCTGAAATTAATCAAGTGAACGAACGAATGGCTT 1235  
DB 1081 TGAATGGGATACAGATTTGCTCTGAAATTAATCAAGTGAACGAACGAATGGCTT 1140  
QY 1236 CACTGTCCGCTTAATCCAGCGTATATTCCTACGCTAACACGAGCGGACACTAAAT 1295  
DB 1141 CACTGTCCGCTTAATCCAGCGTATATTCCTACGCTAACACGAGCGGACACTAAAT 1200  
QY 1296 CGTTTACTTTATGCAATTAATGAAGAAAGCAGATCTTACGAAGGCTTTAAATAGGC 1355  
DB 1201 CGTTTACTTTATGCAATTAATGAAGAAAGCAGATCTTACGAAGGCTTTAAATAGGC 1260  
QY 1356 GAATGTTGAATACGGTCATACCGACGACCAACCAACCACTGTTGAAGTTGTGACAG 1415  
DB 1261 GAATGTTGAATACGGTCATACCGACGACCAACCAACCACTGTTGAAGTTGTGACAG 1320  
QY 1416 TGGGAAACGTTTCATTAATAGATGATGCGATGTGACAGGACCAAGCTTGGCGGAGC 1475  
DB 1321 TGGGAAACGTTTCATTAATAGATGATGCGATGTGACAGGACCAAGCTTGGCGGAGC 1380  
QY 1476 TTCTTTTGTGCTGCGTATCAAAACAGCGACACAGCAATTTATTTGAATATGATGAAC 1535  
DB 1381 TTCTTTTGTGCTGCGTATCAAAACAGCGACACAGCAATTTATTTGAATATGATGAAC 1440  
QY 1536 AACGAAAGCAGAACTTGGGTAAAAACAAACCTGAAGAACTTATTAACCAACGCGC 1595  
DB 1441 AACGAAAGCAGAACTTGGGTAAAAACAAACCTGAAGAACTTATTAACCAACGCGC 1500  
QY 1596 TGATGGAATTAATGATTAACAGGCTTAATACGATACCTATTTATTTGAAGAACTGT 1655  
DB 1501 TGATGGAATTAATGATTAACAGGCTTAATACGATACCTATTTATTTGAAGAACTGT 1560  
QY 1656 AGCTCTGATGATTAATGCTTTGTTAAACAATGGAATGAAATTTGTGCAATGAACAATC 1715  
DB 1561 AGCTCTGATGATTAATGCTTTGTTAAACAATGGAATGAAATTTGTGCAATGAACAATC 1620  
QY 1716 ATATGCAACAAAGAAACCTGATTTACCAAGAAAAGTACCAACCAAGAGTAC 1775  
DB 1621 ATATGCAACAAAGAAACCTGATTTACCAAGAAAAGTACCAACCAAGAGTAC 1680  
QY 1776 CTTACTT 1782  
DB 1681 CTTACTT 1687

RESULT 8  
US-10-912-362-219  
; Sequence 219, Application US/10912362  
; Publication No. US20050043528A1  
; GENERAL INFORMATION:  
; APPLICANT: Choi et al.  
; TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides  
; NUMBER OF SEQUENCES: 497



CORRESPONDENCE ADDRESS:  
ADDRESSER: Human Genome Sciences, Inc.  
STREET: 14200 Shady Grove Road  
CITY: Rockville  
STATE: Maryland  
COUNTRY: USA  
ZIP: 20850  
COMPUTER READABLE FORM:  
MEDIUM TYPE: CD-R  
COMPUTER: Dell Latitude  
OPERATING SYSTEM: Windows 2000  
SOFTWARE: ASCII Text  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/912,362  
FILING DATE: 06-Aug-2004  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 10/206,576  
FILING DATE: 2002-07-29  
APPLICATION NUMBER: US 09/071,035  
FILING DATE: 1998-05-04  
APPLICATION NUMBER: US 60/046,655  
FILING DATE: 1997-05-16  
APPLICATION NUMBER: US 60/044,031  
FILING DATE: 1997-05-06  
APPLICATION NUMBER: US 60/066,009  
FILING DATE: 1997-11-14  
ATTORNEY/AGENT INFORMATION:  
NAME: Hyman, Mark J.  
REGISTRATION NUMBER: 46,789  
REFERENCE/DOCKET NUMBER: PB369PID2  
INFORMATION FOR SEQ ID NO: 219:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1687 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
SEQUENCE DESCRIPTION: SEQ ID NO: 219:  
US-10-912-362-219

Query Match 89.7%; Score 1687; DB 9; Length 1687;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1687; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 96 AGAAGAAAATGGGGAGAGCGGACAGCTCGTGATTCACAAAAAGAAAATGACGATTTACC 155  
DB 1 AGAAGAAAATGGGGAGAGCGGACAGCTCGTGATTCACAAAAAGAAAATGACGATTTACC 60  
QY 156 AGATCCGCTTATTCAAAATAGCGGGAGAAAAGAAATGAGCGAGTTGATTAATATCAAGACT 215  
DB 61 AGATCCGCTTATTCAAAATAGCGGGAGAAAAGAAATGAGCGAGTTGATTAATATCAAGACT 120  
QY 216 GGCAGATGTGACGTTTATGATTTATTAACGTGACGAACGAATTTTACGAGCAACGAGCGGC 275  
DB 121 GGCAGATGTGACGTTTATGATTTATTAACGTGACGAACGAATTTTACGAGCAACGAGCGGC 180  
QY 216 AGGCGCAAGCGTTGATGACGCTTAACACAGCTGTCCAAAGTTTAACTCTCGGAAACTGT 335  
DB 276 AGGCGCAAGCGTTGATGACGCTTAACACAGCTGTCCAAAGTTTAACTCTCGGAAACTGT 335  
QY 181 AGGCGCAAGCGTTGATGACGCTTAACCAAGCTGTCCAAAGTTTAACTCTCGGAAACTGT 240  
DB 336 TGCTCAAGGAACCAACGATGCAAAATGGGAATGTCACTGTTCACTTAACCTAAACAAACAA 395  
QY 241 TGCTCAAGGAACCAACGATGCAAAATGGGAATGTCACTGTTCACTTAACCTAAACAAACAA 300  
DB 396 TGGTAAAGATGACGTGTATACCATTAAGAGAAACCAAAAGAGGTGTAGTGTGCTTAC 455  
QY 301 TGGTAAAGATGACGTGTATACCATTAAGAGAAACCAAAAGAGGTGTAGTGTGCTTAC 360  
DB 456 GAATATGCTGTGCGGCTTCCAGTTTACGAATGATCAAGCAAAACGATGCTTCTATTA 515  
QY 361 GAATATGCTGTGCGGCTTCCAGTTTACGAATGATCAAGCAAAACGATGCTTCTATTA 420  
DB 516 ATATGAAACAGAAATTAAGCGTTGTTCATATTTTCTAAAAATGTGTGACCAATGA 575

DB 421 ATATGAAACAGAAATTAAGCGTTGTTCATATTTATCTAAAAATGTGTAGCCAAATGA 480  
QY 576 TGGTATGTTTACATGTAAGAAAAGTGTGAACCTGCTGAAAATGAAGATTTAAATGGCGCAGA 635  
DB 481 TGGTATGTTTACATGTAAGAAAAGTGTGAACCTGCTGAAAATGAAGATTTAAATGGCGCAGA 540  
QY 636 ATTGCTTAATTTCTAAAGCGAAAGGCTCACAGGCAAGTAAATATATCAAGAGTCA 695  
DB 541 ATTGCTTAATTTCTAAAGCGAAAGGCTCACAGGCAAGTAAATATATCAAGAGTCA 600  
QY 696 AGATGATTAATATATCATGACCAACGATTAAGAAACAAAGCAAAACGCTTTATTAAGTGGGA 755  
DB 601 AGATGATTAATATATCATGACCAACGATTAAGAAACAAAGCAAAACGCTTTATTAAGTGGGA 660  
QY 756 AAGTTATGAATTTGGCGAAAATGATTTTACAGAACAGAGAAATGGAACGGGGAATTAAC 815  
DB 661 AAGTTATGAATTTGGCGAAAATGATTTTACAGAACAGAGAAATGGAACGGGGAATTAAC 720  
QY 816 AGTTAAATATCTTGAGTTGGTGTGATATTTTAAAGAAAGTAAAGCTCCAAATTAATGC 875  
DB 721 AGTTAAATATCTTGAGTTGGTGTGATATTTTAAAGAAAGTAAAGCTCCAAATTAATGC 780  
QY 876 AGAATTAATTTGAAAATCAACCAAAACCAATTTTCAATTGAAGCAAAATCAAAACACC 935  
DB 781 AGAATTAATTTGAAAATCAACCAAAACCAATTTTCAATTGAAGCAAAATCAAAACACC 840  
QY 936 TGTGAAAAAACAGTCAAAAATGATACCTCTTAAAGTTGATTAACCAACCAAGCTTGA 995  
DB 841 TGTGAAAAAACAGTCAAAAATGATACCTCTTAAAGTTGATTAACCAACCAAGCTTGA 900  
QY 996 TGGTAAAGATGCGCAATTTGGGAAAATTAATTAATCAAAATTTCTGTAATATTCATT 1055  
DB 901 TGGTAAAGATGCGCAATTTGGGAAAATTAATTAATCAAAATTTCTGTAATATTCATT 960  
QY 1056 GGGGATTCAGACAAAGAAAGGCGCTAATTAATATAGCTCAAAATTTAGTTGATTA 1115  
DB 961 GGGGATTCAGACAAAGAAAGGCGCTAATTAATATAGCTCAAAATTTAGTTGATTA 1020  
QY 1116 ACATGATGACGCTTAACTTTTGAATAGCTGCTTGTGAAGTATGCTTATAGCTTATA 1175  
DB 1021 ACATGATGACGCTTAACTTTTGAATAGCTGCTTGTGAAGTATGCTTATAGCTTATA 1080  
QY 1176 TGATGGGATACAGGATTTGCTCTGAATAATTAATCAAGTGAACCAAGCAAAATGGGCT 1235  
DB 1081 TGATGGGATACAGGATTTGCTCTGAATAATTAATCAAGTGAACCAAGCAAAATGGGCT 1140  
QY 1236 CACTGTGCGCGTTAATCCAGCGTATATTCCTAAGCTTAACCAAGCGGCGACCTAAAT 1295  
DB 1141 CACTGTGCGCGTTAATCCAGCGTATATTCCTAAGCTTAACCAAGCGGCGACCTAAAT 1200  
QY 1296 CGTTTACTTATGCAATTTTAAATGAAGAAAGCAGATCCTTACGAAGGCTTTAAATGAAGC 1355  
DB 1201 CGTTTACTTATGCAATTTTAAATGAAGAAAGCAGATCCTTACGAAGGCTTTAAATGAAGC 1260  
QY 1356 GAAATGTGATACGGTCAATACGAGACCAACCAACCAACTGTTGAAGTGTGACAG 1415  
DB 1261 GAAATGTGATACGGTCAATACGAGACCAACCAACCAACTGTTGAAGTGTGACAG 1320  
QY 1416 TGGGAAGCGTTTCATTAAGTGTGATGCGAGTGTGACAGGACCAAGGCTTGGCGGAGC 1475  
DB 1321 TGGGAAGCGTTTCATTAAGTGTGATGCGAGTGTGACAGGACCAAGGCTTGGCGGAGC 1380  
QY 1476 TTCTCTTGTGCTGCGGTGATCAAAACAGGACCAAGCAATTTTGAAGAAATGATGAAC 1535  
DB 1381 TTCTCTTGTGCTGCGGTGATCAAAACAGGACCAAGCAATTTTGAAGAAATGATGAAC 1440  
QY 1536 AACGAAGAGCAACTGGGAGTGAAGAAACAAAGCTGAAGCACTACTTTTACAAACAGGC 1595  
DB 1441 AACGAAGAGCAACTGGGAGTGAAGAAACAAAGCTGAAGCACTACTTTTACAAACAGGC 1500  
QY 1596 TGATGATTAATGTTGATATCAAGGGCTTAAATACGATCTTATTTTGAAGAAACCTGT 1655  
DB 1501 TGATGATTAATGTTGATATCAAGGGCTTAAATACGATCTTATTTTGAAGAAACCTGT 1560



QY 1656 AGCTCTGATGATTAATGCTGTTTAAACAATGGATTTGATTTGCTGATTAACAATC 1715  
| | | | |  
Db 1561 AGCTCTGATGATTAATGCTGTTTAAACAATGGATTTGATTTGCTGATTAACAATC 1620  
| | | | |  
QY 1716 ATATGGCAACAAGAAACCTAGTTTACCAGAAAAGTACCAACAACAAGAGTAC 1775  
| | | | |  
Db 1621 ATATGGCAACAAGAAACCTAGTTTACCAGAAAAGTACCAACAACAAGAGTAC 1680  
| | | | |  
QY 1776 CTTACT 1782  
| | | | |  
Db 1681 CTTACT 1687  
| | | | |  
  
RESULT 9  
US-10-661-809-18.  
; Sequence 18, Application US/10661809  
; Publication No. US20040101919A1  
; GENERAL INFORMATION:  
; APPLICANT: HOOK, Magnus  
; TITLE OF INVENTION: BIOINFORMATIC METHOD FOR IDENTIFYING SURFACE-ANCHORED PROTEINS FR  
; FILE REFERENCE: POSITIVE BACTERIA  
; CURRENT APPLICATION NUMBER: US/10/661,809  
; CURRENT FILING DATE: 2003-09-15  
; PRIOR APPLICATION NUMBER: 60/410303  
; NUMBER OF SEQ ID NOS: 24  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 18  
; LENGTH: 1878  
; TYPE: DNA  
; ORGANISM: Staphylococcus epidermidis  
US-10-661-809-18  
  
Query Match 53.9%; Score 1014.6; DB 8; Length 1878;  
Best Local Similarity 71.9%; Pred. No. 6, 2e-194;  
Matches 1342; Conservative 0; Mismatches 519; Indels 6; Gaps 1;  
  
QY 1 ATGAAGCAATTAATAAAAGTTGGTGTACACCGTTAGTACTTGTACTAATTTGGCACTT 60  
| | | | |  
Db 1 ATGAAGCAATTAATAAAAGTTGGTGTACACCGTTAGTACTTGTACTAATTTGGCACTT 60  
| | | | |  
QY 61 TTCAAGGTATTAAGGACACACCTGATTTGAGAAAGAGGAGAGCCGACAG 120  
| | | | |  
Db 61 CTCAACAAAGCTTGCGGCAAAAAGGTTGACAGAGAGACAG-----CAGCTCA 114  
| | | | |  
QY 121 CTCGTGATTCACAAAAGAAATGACGATTTACAGATCCGCTTATTTCAAAATAGCGG 180  
| | | | |  
Db 115 GTCAATCTTCAATAAAGAAATGACGATTTACCGATCTTTAATCCAAACAGCGG 174  
| | | | |  
QY 181 AAAGAAATGAGCGATTTGATTAATATCAAGACTGGCAGATGTGACGTTAGTAT 240  
| | | | |  
Db 175 AAAGAAATGAGCGATTTGATTAATATCAAGACTGGCAGATGTGACGTTAGTAT 234  
| | | | |  
QY 241 AACGAGAGAGCAATTTTAAGCAAGAGGAGGAGCGCAAGCGTTGATTCAGACTAA 300  
| | | | |  
Db 235 AACGAGAGAGCAATTTTAAGCAAGAGGAGGAGCGCGTCCGTGATGACAGAAA 294  
| | | | |  
QY 301 CAAGCTGTCAAAGTTTAATCTCTGGGAAACCTGTGCTCAAGAAACACCGATGCAAT 360  
| | | | |  
Db 295 CAAGCAGTCAGATCTTTGATCTCTGTGACACCAAGTTGCTTCAAGAAACAGATGCTAT 354  
| | | | |  
QY 361 GGAATGTCACTGTTCACTTACCTTAAACAAATGTTAAAGATGACGTTATACAT 420  
| | | | |  
Db 355 GGAATGTCACTTATCTTAACTTAAACAAATGTTAAAGATGACGTTATACAT 414  
| | | | |  
QY 421 AAAGAAAGACCAAAAGAGGAGTGTAGTGTCTTACGAATATGAGTGTGCGCTTCCAGT 480  
| | | | |  
Db 415 AAAGAAAGACCAAAAGAGGAGTGTAGTGTCTTACGAATATGAGTGTGCGCTTCCAGT 474  
| | | | |  
QY 481 TACGAATGATCAACCAACAGATGTTCTTAAATATGAAACAGAAATTTAGCGTT 540  
| | | | |

Db 475 TATGAGATGATCAACCAAGCAGATGCGCTTATTAATAATAGGGACAGAAAGACTAGTACT 534  
| | | | |  
QY 541 GTTCAATATTATCTTAAATATGTTGTAGCCAAATGATGTGATTTACATGTGAAAAAGTA 600  
| | | | |  
Db 535 ATTCATCTCTACCTTAAATATGATGCGGTAAATGATGAAACGTTGAAATTCAAAAATC 594  
| | | | |  
QY 601 GGAATGCGGAAAAATGAAGATTTAAATGGCGGAGAAATTTGTTATTTCTTAAAGCGAAGC 660  
| | | | |  
Db 595 GGTACTGCCGAAACCAAGCAGCTTAAATGAGAGCAAGATTTATTTCTTAAAGAAAGAG 654  
| | | | |  
QY 661 TCACAGAGCAGATTAATATATATCCAGAGAGTCAAGATGATATATATCATGACAGAG 720  
| | | | |  
Db 655 ACACCAAGCGTCAAAAAATATATCCAGAGTGTCAAGATGATTTGTACCTTGGACAACT 714  
| | | | |  
QY 721 GATTAAGAAACAGCAAAACGCTTTATTTACTGAGAAAAGTTATGAATTTGGCGAAATGAT 780  
| | | | |  
Db 715 GATCAAAACCAAGCAACATTTCAATCTGATCTATTTGACATCGGCAACATGAC 774  
| | | | |  
QY 781 TTCAAGAGCAGAGATGGAACGGGAGAAATTAACGTTAAATAATCTGAGTTGGTTGG 840  
| | | | |  
Db 775 TTGCGGAGGCAATCTATTAAGAAAAGGCCAGTTGATGTTAATCAATTTAGAAAGTTGAAAA 834  
| | | | |  
QY 841 TATATTTTGAAGAGTAAAGCTCCAAATATATGAGAAATTAATGAATCAACAA 900  
| | | | |  
Db 835 TATATTTTGAAGAGTAAAGCTCCGATATATGCGAAATGATTAAGAAACAAATC 894  
| | | | |  
QY 901 ACAATTTTCAATTTGAAAGCAATCAACACCTGTTGAAAACAGTCAAAATGAT 960  
| | | | |  
Db 895 AGCCTTTGAGATCTGCGCAATGACCAACACAGATGAAAGCCATCAAAATGAT 954  
| | | | |  
QY 961 ACCCTAAAGTTGATTAACCAACACCAAGCTAGATGTTAAAGTGTGCAATTTGGCGAA 1020  
| | | | |  
Db 955 AGCTTAAAGTTGATTAACCAACACCTCAATTAATGAAAGATGTCGAATCGGTGA 1014  
| | | | |  
QY 1021 AAAATTAATATCAATTTCTGTAATATTTCAATTTGGGATTCAGACAAAGGCGAC 1080  
| | | | |  
Db 1015 AAAATTAATATGATTTCTGTAATATTTCAATTTGGGATTCAGACAAAGGCGAC 1074  
| | | | |  
QY 1081 GCTAATTAATAGCTCAATTTCAATTTAGTTAATCAATGATGACGCTTAACTTTGAT 1140  
| | | | |  
Db 1075 CAACCAAGTACACCAACATTCACATTCATGATCTGATGACGCTTAACTTTGAT 1134  
| | | | |  
QY 1141 AACGTGCTTCGAGAGATATGCTTATGCTTATGATTTGGGATGACGTGTTGCTCT 1200  
| | | | |  
Db 1135 AATGATTTCTGAGAAAGTATGCTTATGCTTATGATTTGGGATGACGTGTTGCTCT 1194  
| | | | |  
QY 1201 GAAATTAATCAAGTACATGAACCAAGCAATGCTTCACTGTGCGCTTATTCAGCGTAT 1260  
| | | | |  
Db 1195 GTAATTAATTTCTGATCTGAGAAACAGACGATTTCAAGTTTCAAGTTGATTCGAATAT 1254  
| | | | |  
QY 1261 ATTCTTACGCTTACACAGGCGGCACTTAAATTCGTTTACTTTATGCAATTTAAATGA 1320  
| | | | |  
Db 1255 ATTCTTATTAATCTCTGCGGCTTCAATGAAATTCGTTTACTATATGATTTGAACGA 1314  
| | | | |  
QY 1321 AAAGCAATCTTACGAAAGGCTTTAAATGAGGGAATGTTGATTAAGGCTATCCGAC 1380  
| | | | |  
Db 1315 AAAGCAATCTTACGAAAGGCTTTTAAACAGAGAAATGCTATTAAGGCTATCCAAAT 1374  
| | | | |  
QY 1381 GACCAACACCAACCACTGTTGAAGTTGTGACAGTGGGAAACGTTTCAATTAAGTCAAT 1440  
| | | | |  
Db 1375 GATCAACACCAACCGTCAAGTGTGATGTTGTTGAGGCAACGATTTGTTAAATGAAT 1434  
| | | | |  
QY 1441 GCGATGTGACAGGACCAAGCTTTGGCGGAGCTTCTTGTGTGCTGCTGATCAAAAC 1500  
| | | | |  
Db 1435 GGTGACGTTTACATCAACCAACCTTCTGAGAGCAATTCGTGCTGATATCAAGAT 1494  
| | | | |  
QY 1501 AGCGACAGCAAAATTTTGAATAATGATGAACCAAGAAAGCAGCACTTGGGTGAAA 1560  
| | | | |  
Db 1495 AGTGAACAGGAAATTTATGATGACCAACATTCAGCAAAAGCGGTGACGTGGTATG 1554  
| | | | |  
QY 1561 ACAAAGCTGAACCACTTATTTTCAACCAACGCGTGTGATTAATGTTGATATCAAGAG 1620  
| | | | |  
Db 1555 GCGAAAGATCAAGCAAGGTTTATCAACCAAGTAAAGGTTTAAATGATGTGACAGAT 1614  
| | | | |

```
QY 1621 CTTAATACGTAACCTATTTAGAGAACTGTAAGCTCGATGATTAATGCTTGTGA 1680
Db 1615 CTAATAATATGCACTACTATCTGAAAGAAAGCAAGGCGCCAGAAAAATATTTCCATTA 1674
QY 1681 ACAATCGGATTTGAAATTTGTGTCATATGCAATCATATGAGCAACAGAAAACTAGTT 1740
Db 1675 ACAACCGGTATGCACTTATCTATCGATGAAACAATCTTATGTAAACAGGAGCAAGTTGANT 1734
QY 1741 TCACCAAGAAAAAGTACCAACCAACCAAGGTAACCTTACCTTCAACAGGTGGCAAGGA 1800
Db 1735 TCTCTGAAAAAATATCCAAATTAACCAAGGTAACCTTCTTCAACAGGCGGTAAAGGA 1794
QY 1801 ATCTAGTTTACTTGTGAAGTGGCCAGTCTTACTATTTGCAAGAGTCTACTTGTCT 1860
Db 1795 ATCTATGATATATGCGTGCAGAGTATGCTTCTTACTATGCTGAGACTGTTTGTCT 1854
QY 1861 AGACGTA 1867
Db 1855 AGACGCA 1861

RESULT 10
US-10-333-002-28
; Sequence 28, Application US/10333002
; Publication No. US20040071729A1
; GENERAL INFORMATION:
; APPLICANT: Adderson, Elisabeth
; TITLE OF INVENTION: Group B Streptococcus Polypeptides Nucleic Acids and Therapeutic
; FILE REFERENCE: 2511-1-001 (Sf-0039)
; CURRENT APPLICATION NUMBER: US/10/333,002
; PRIOR FILING DATE: 2003-07-08
; PRIOR APPLICATION NUMBER: PCT/US01/24795
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 09/534,341
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 28
; TYPE: DNA
; LENGTH: 2199
; ORGANISM: Enterococcus faecalis
US-10-333-002-28

Query Match 31.6%; Score 595; DB 8; Length 2199;
Best Local Similarity 100.0%; Pred. No. 2.2e-109;
Matches 595; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1287 ACTAAATTCGTTTCTTATGATTTAATGAAAAAGCAGATCCTAGAAAGGCTTTAA 1346
Db 1 ACTAAATTCGTTTCTTATGATTTAATGAAAAAGCAGATCCTAGAAAGGCTTTAA 60
QY 1347 AAATGAGCGAATGTTGATATCGTCAATCCGACGACCAACCAACCACTGTTGAAGT 1406
Db 61 AAATGAGCGAATGTTGATATCGTCAATCCGACGACCAACCAACCACTGTTGAAGT 120
QY 1407 TGTGACAGTGGGAAAAGCTTTCATTAAGTCAATGCGCATGTGACAGGACACACGCTT 1466
Db 121 TGTGACAGTGGGAAAAGCTTTCATTAAGTCAATGCGCATGTGACAGGACACACGCTT 180
QY 1467 GCGGGAGGCTTCCTTTGTGTCGCGTATCAAAAACAGGACACAGCAAAATTTATGAAAT 1526
Db 181 GCGGGAGGCTTCCTTTGTGTCGCGTATCAAAAACAGGACACAGCAAAATTTATGAAAT 240
QY 1527 CGATGAAACAAAGAAAGCAACCTTGGGTGAAAACAAAAGCTGAAGCACTACTTTTAC 1586
Db 241 CGATGAAACAAAGAAAGCAACCTTGGGTGAAAACAAAAGCTGAAGCACTACTTTTAC 300
QY 1587 AACCAACGCTGATGATTTGTTGATATCAAGGGCTTTAAATAGCGTATCTATTTATTA 1646
Db 301 AACCAACGCTGATGATTTGTTGATATCAAGGGCTTTAAATAGCGTATCTATTTATTA 360
```

```
QY 1647 AGAACTGTAAGTCTGATGATTTATGTTTAAACAAATGGATTAATTTGTGTCAA 1706
Db 361 AGAACTGTAAGTCTGATGATTTATGTTTAAACAAATGGATTAATTTGTGTCAA 420
QY 1707 TGAACATCATATGACACAAAGAAACCTAGTTTCAACAGAAAAAGTACCAACAAACA 1766
Db 421 TGAACATCATATGACACAAAGAAACCTAGTTTCAACAGAAAAAGTACCAACAAACA 480
QY 1767 CAAAGTACCTTACTTCAACAGGTGGCAAGAAATCTACCTTCTTAGGAAGTGGCC 1826
Db 481 CAAAGTACCTTACTTCAACAGGTGGCAAGAAATCTACCTTCTTAGGAAGTGGCC 540
QY 1827 AGCTTGTCTACTTATTTGACAGAGTCTACTTTGCTAGAGTGAAGAAATGCT 1881
Db 541 AGCTTGTCTACTTATTTGACAGAGTCTACTTTGCTAGAGTGAAGAAATGCT 595

RESULT 11
US-10-425-115-120013/C
; Sequence 120013, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; PRIOR FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 120013
; LENGTH: 1062
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(1062)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_40935C.1
US-10-425-115-120013

Query Match 3.6%; Score 67; DB 9; Length 1062;
Best Local Similarity 44.0%; Pred. No. 0.0037;
Matches 283; Conservative 0; Mismatches 360; Indels 0; Gaps 0;

QY 593 AAAAAGTGAAGTCTGTAATGAAGATTAATGCGCGCAGAAATTTGTTATTTCTAAA 652
Db 970 AAAAAGTGAAGTCTGTAATGAAGATTAATGCGCGCAGAAATTTGTTATTTCTAAA 911
QY 653 GCGAAGCTCACAGGACAGTAAATATATCAAGAGTCAAAGTGAATATATATCAT 712
Db 910 AAAAAGTGAAGTCTGTAATGAAGATTAATGCGCGCAGAAATTTGTTATTTCTAAA 851
QY 713 GGAACAAGGATTAAGAAACAAAGCAACGCTTATTTACTGGAAGAAATTAATGAAATTTGGC 772
Db 850 AAAAAGTGAAGTCTGTAATGAAGATTAATGCGCGCAGAAATTTGTTATTTCTAAA 791
QY 773 AAAATGATTTCAAGAGCAGAAATGAAACGCGAATTTACAGTTAAATCTTGAAG 832
Db 790 AAAAAGTGAAGTCTGTAATGAAGATTAATGCGCGCAGAAATTTGTTATTTCTAAA 731
QY 833 TTGCTTCTGATATTTTGAAGAGTAAAGCTCCAAATATATCAAGATTTATGAAATC 892
Db 730 AAAAAGTGAAGTCTGTAATGAAGATTAATGCGCGCAGAAATTTGTTATTTCTAAA 671
QY 893 AAAAAGTGAAGTCTGTAATGAAGATTAATGCGCGCAGAAATTTGTTATTTCTAAA 952
Db 670 AAAAAGTGAAGTCTGTAATGAAGATTAATGCGCGCAGAAATTTGTTATTTCTAAA 611
```

Qy 953 AAAATGATCTCTAAAGTGTAAACACACAGCTTAGGTGTAAGATGTGGCAA 1012  
Db 610 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 551  
Qy 1013 TTGGCGAAAAATTAATATCAATTTCTGTAAATATTCATTGGGATTCAGACAAAG 1072  
Db 550 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 491  
Qy 1073 AAGGCGACCTAATTAATACGTCAATTCATTTAGTGTAAACGTGACGCTTAA 1132  
Db 490 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 431  
Qy 1133 CTTTGATATGAGCTTCTGAGAGATGCTTATGCTTATATGATGGGATTCAGTGA 1192  
Db 430 GAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 371  
Qy 1193 TTGCTCTGAAAAATTAATCAAGTGAACAGCAAAATGGCTT 1235  
Db 370 ATATATATGACAAAAACAGCAACAAAAACAAAGCGGCAT 328

## RESULT 12

US-10-021-323-9336  
; Sequence 9336, Application US/10021323  
; Publication No. US20040123340A1  
; GENERAL INFORMATION:  
; APPLICANT: Deikman, J111  
; APPLICANT: Feng, Paul C.C.  
; APPLICANT: Fincher, Karen L.  
; APPLICANT: Ziegler, Todd E.  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(52274)B  
; CURRENT APPLICATION NUMBER: US/10/021,323  
; PRIOR FILING DATE: 2001-12-12  
; PRIOR APPLICATION NUMBER: US 60/255, 619  
; PRIOR FILING DATE: 2000-12-14  
; NUMBER OF SEQ ID NOS: 17880  
; SEQ ID NO 9336  
; LENGTH: 627  
; TYPE: DNA  
; ORGANISM: *Gosypium hirsutum*  
; FEATURE:  
; OTHER INFORMATION: Clone ID: LIB3828-024-Q6-K6-C2  
US-10-021-323-9336

Query Match 3.5%; Score 65; DB 8; Length 627;  
Best Local Similarity 46.1%; Pred. No. 0.0075;  
Matches 218; Conservative 0; Mismatches 255; Indels 0; Gaps 0;

Qy 649 AAAACCGAAGCTCACCGGACAGTAAATATTCAGAGGTCAAAGATGATTTAT 708  
Db 124 AAAAAAAAAAGAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 183  
Qy 709 ACATGACACGATTAAGAACACAAACGCTTATATGCGAAAGTATGAAAT 768  
Db 184 AAAAAAAAAAAAAAAAAAAAAAAAAAGAAAGAAAAAAGAAAAAAGAAAAA 243  
Qy 769 GCGCAAAATGATTTACAGAGACAGAAATGAGCGGAAATTAACAGTTAAATCTT 828  
Db 244 GAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 303  
Qy 829 GAGTTGTTCTATATTTTGAAGAAGTAAAGCTCCAAATTAATGCAATTAATGAA 888  
Db 304 AAGGAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 363  
Qy 889 AATCAACAAAAACACATTTTCAATTGACAAACAAATCAACACCTGTTGAAAAACA 948  
Db 364 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 423  
Qy 949 GTCAAAATGATCTTAAGTGTAAACACACACCAAGCTTGAAGTGAAGATGTG 1008  
Db 424 AAAAAAAAAAGAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 483

Qy 1009 GCAATTTGGCGAAAAATTAATATCAAAATTTCTGTAATATTCATTTGGGATTCAGAC 1068  
Db 484 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 543  
Qy 1069 AAGAAAGCGACGCTAATTAATATGCTCAATTTCAATTTAGTGTAAACATGA 1121  
Db 544 AAAAAAAAAAGAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 596

## RESULT 13

US-10-172-502-1  
; Sequence 1, Application US/10172502  
; Publication No. US20030185833A1  
; GENERAL INFORMATION:  
; APPLICANT: FOSTER, Timothy et al.  
; TITLE OF INVENTION: CROSS-REACTIVE MONOCLONAL AND POLYCLONAL ANTIBODIES.  
; FILE REFERENCE: P07263US01/BAS  
; CURRENT APPLICATION NUMBER: US/10/172,502  
; PRIOR FILING DATE: 2002-06-17  
; PRIOR APPLICATION NUMBER: US 60/298,098  
; PRIOR FILING DATE: 2001-06-15  
; NUMBER OF SEQ ID NOS: 29  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 1  
; LENGTH: 6609  
; TYPE: DNA  
; ORGANISM: *Staphylococcus epidermidis*  
US-10-172-502-1

Query Match 3.4%; Score 64.6; DB 7; Length 6609;  
Best Local Similarity 46.9%; Pred. No. 0.023;  
Matches 202; Conservative 0; Mismatches 229; Indels 0; Gaps 0;

Qy 545 ATATTTATCTTAATATGATGTAGCCATGATGTAGTTACATGTGAAGAAAGTAGAA 604  
Db 5549 ATACGAGCGACAGATGAAGAACAAATGCTGCAATAGTACAGTTGAAAAAGATTAA 5608  
Qy 605 CTGCTGAAAAATGAAGATTTAAATGGCGCAATTTGTTATTTCTTAAGCGAAGCTCAG 664  
Db 5609 TTAAAGCTTAACAAATATGCTGTGCAAGTCAATGCTGATGTGCTATTTATTTGCG 5668  
Qy 665 CAGGACAGTAAATATATTCAGAGGTCAAGATGATGATTAATATCATGACACGAGATA 724  
Db 5669 ATGATGGAAGAAACAAATTTGCTGAATTCGAATCTGTTATTAATTAAGCACTGCGC 5728  
Qy 725 AAGAACAGCAAAACGCTTTATTTACTGGAAGAAAGTTATGAATTTGCGAAATTTGA 784  
Db 5729 GAGAACATTAACAATTTATTTCAAGATTAAGAACAGCAATTTGAAGCAATTTCAAG 5788  
Qy 785 CAGAACAGAGATGAACGGAAGATTAACAGTTAAATTTTGAAGTTGTTGTTGTTATA 844  
Db 5789 CAACGTGAAGAAAGAAATGATTTTATGACAGTTCAAAACATTAATGACACTGCTA 5848  
Qy 845 TTTTGAAGAAATTAAGCTCAATTAATGACAAATTAATGAATTAATGAATTAATGA 904  
Db 5849 TTGCAAAATGATTAAGATGCTAGCAATGCAAGATTTGATTAACACAGCACTTAATTC 5908  
Qy 905 CATTTACATTTGAACCAACATCAACACCTGTTGAAGAAACAGTCAAAATGATACCT 964  
Db 5909 TACAAACAATCATGATTTAGCGTACATCTTAATTAAGCAAGATGCTGAAAAACGA 5968  
Qy 965 CTAAGTTGAT 975  
Db 5969 TTAATGATGAT 5979

## RESULT 14

US-11-020-509-1  
; Sequence 1, Application US/11020509  
; Publication No. US20050106648A1  
; GENERAL INFORMATION:  
; APPLICANT: FOSTER, Timothy et al.

;; TITLE OF INVENTION: CROSS-REACTIVE MONOCLONAL AND POLYCLONAL ANTIBODIES. . .  
;; FILE REFERENCE: P07263US02/BAS  
;; CURRENT APPLICATION NUMBER: US/11/020,509  
;; CURRENT FILING DATE: 2004-12-27  
;; PRIOR APPLICATION NUMBER: US 10/172,502  
;; PRIOR FILING DATE: 2002-06-17  
;; PRIOR APPLICATION NUMBER: US 60/298,098  
;; PRIOR FILING DATE: 2001-06-15  
;; NUMBER OF SEQ ID NOS: 29  
;; SOFTWARE: Patent version 3.1  
;; SEQ ID NO 1  
;; LENGTH: 6609  
;; TYPE: DNA  
;; ORGANISM: Staphylococcus epidermidis  
US-11-020-509-1

Query Match 3.4%; Score 64.6; DB 13; Length 6609;  
Best Local Similarity 46.9%; Pred. No. 0.023;  
Matches 202; Conservative 0; Mismatches 229; Indels 0; Gaps 0;

QY 545 ATATTATCTTAAAGTGTGAGCAATGATGTTTACATGTAAGAAAAGTAGAA 604  
DB 5549 ATACGAAGCGACACATGAAAGAAACAATGCTGCAATGATACAGTTGAAAAGAGTTAA 5608  
QY 605 CTGCTGAAATGAAGATTAAATGCGCAGAAATTTGTTATTTCTAAAAGCGAGCTCAC 664  
DB 5609 TTAAGCTTAACACAAATGCTGCTGACGATGATGCTGATGCGCATTTTATTGCG 5668  
QY 665 CAGGACAGTAAATATATTCAGAGAGTCAAAAGATGATTATATCATGACACAGGATA 724  
DB 5669 ATGATGGGAAAACGAAATTCCTGGAATCGAACCTGTTATTTAAATAAAAGCACTGCGC 5728  
QY 725 AAGACAGCAAAAGCGTTTATTACGTGGGAAAAGTTAGAAATTCGCGAAAATGATTTCA 784  
DB 5729 GAGAACATTTCACATTTATTCAACGATAGAAAACAGCAATGAGGAAATGTTCAAG 5788  
QY 785 CAGAACAGAGAAATGAAACGGAGATTAAACAGTTAAATCTTGAAGTTGGTTGCTATA 844  
DB 5789 CAACGATGAGAAAGAAATAGTATTTTAGACAGTTTCAAAAACATTTATGACAGTGGTA 5848  
QY 845 TTTTGAAGAAATAAAGCTCCAATATATGCAATTTAATTGAAAATCAAAACAAACAC 904  
DB 5849 TTGCAAAATTTGATCAAGATCTGACCAATGCAAAAGTTGATAAAACAGCAATTTAAATC 5908  
QY 905 CATTACATTGAACCAACATCAACACCTGTTGAAAAAAGCAAAAATGATACCT 964  
DB 5909 TACAAACATCATGATTTAGACGTACATCTATTAAAAAGCCAAATGCTGAAAAACGA 5968  
QY 965 CTAAAGTTGAT 975  
DB 5969 TTAATGATGAT 5979

RESULT 15  
US-10-425-115-123086/c  
;; Sequence 123086, Application US/10425115  
;; Publication No. US20040214272A1  
;; GENERAL INFORMATION:  
;; APPLICANT: La Rosa, Thomas J.  
;; APPLICANT: Kovalic, David K.  
;; APPLICANT: Zhou, Yihua  
;; APPLICANT: Cao, Yongwei  
;; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
;; FILE REFERENCE: 38-21(53222)B  
;; CURRENT APPLICATION NUMBER: US/10/425,115  
;; CURRENT FILING DATE: 2003-04-28  
;; NUMBER OF SEQ ID NOS: 369326  
;; SEQ ID NO 123086  
;; LENGTH: 1029  
;; TYPE: DNA  
;; ORGANISM: Zea mays  
;; FEATURE:

;; NAME/KEY: unsure  
;; LOCATION: (1)-(1029)  
;; OTHER INFORMATION: unsure at all n locations  
;; FEATURE:  
;; OTHER INFORMATION: Clone ID: MRT4577\_43736C.1  
US-10-425-115-123086

Query Match 3.4%; Score 64.4; DB 9; Length 1029;  
Best Local Similarity 45.0%; Pred. No. 0.012;  
Matches 239; Conservative 0; Mismatches 292; Indels 0; Gaps 0;

QY 612 AAATGAGATTAAATGCGCAGAAATTTGTTATTTCTAAAAGCGAAGCTCACAGGAC 671  
DB 1011 AAAAAAAAAAGAGAAAAAATTTAAATTTACAAAAAAGAAAAAAGAAAAA 952  
QY 672 AGTAAATATATCCAGAGTCAAGATGATTTATATACATGACACAGATTAAGACA 731  
DB 951 AAAAAAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 892  
QY 732 AGCAAAACGCTTTTATTAATGGAAGATTAATGAAATTTGCGAAATTTGACAGAAC 791  
DB 891 CAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 832  
QY 792 AGAGATGAAACGGAGATTACAGTTAAATCTTGAAGTTGGTTGCTATATTTTGA 851  
DB 831 AAAAAAAAAAAGAAAAAATGATTAATTAATTAATTAATTAATTAATTAATTAATTA 772  
QY 852 AGAATTAAGCTCAATTAATGCAATTTAATTAATTAATTAATTAATTAATTAATTA 911  
DB 771 AAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 712  
QY 912 AATTGAGCAAAACATCAACACCTGTTGAAAAAAGCTCAAAAATGATCTTAAAGT 971  
DB 711 AAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 652  
QY 972 TGATTAACCAACCAAGCTTATGATGTAAGATGTGCAATTTGCGGAAAAATTAATA 1031  
DB 651 AAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 592  
QY 1032 TCAATTTCTGTAATTTATTCATTGGGATTTGACAGCAAAAGGCGACGCTTAATA 1091  
DB 591 AAACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 532  
QY 1092 CGTCAATTTCAATTTAGTTGATTAACATGATGCAAGCTTTAATTTGATAA 1142  
DB 531 AAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 481

Search completed: July 6, 2006, 00:48:21  
Job time : 1680 secs



```
QY      852 AGAAGTAAAGCTCCAAATTAATGCAATTAATTGAAATCAACAAAACACATTAC 911
      2123 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2182
QY      912 AATTGAGCAAAACAATCAACACCTGTTGAAAAAAGCTCAAAATGATCTCTAAAGT 971
      2183 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2242
QY      972 TGTATTAACACACCAAGCTTATGATGTTAAGATGTGCAATTGGCGAAAAATTAAATA 1031
      2243 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2302
QY      1032 TCAATTTCTGTAATTAATTTCCATTGGGGATTGCGACAAAGAGCGACCTAATTAATA 1091
      2303 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2362
QY      1092 CGTCAAAATTCATTTAGTTGATTAACA 1118
      2363 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2389
Db
```

```
RESULT 2
US-10-471-571A-859
; Sequence 859, Application US/10471571A
; Publication No. US20060115490A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SPA
; TITLE OF INVENTION: STAPHYLOCOCCUS AUREUS PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE: P026927WO
; CURRENT APPLICATION NUMBER: US/10/471,571A
; PRIOR FILING DATE: 2003-09-12
; PRIOR APPLICATION NUMBER: GB-0107661.1
; NUMBER OF SEQ ID NOS: 5642
; SOFTWARE: SeqMan9, version 1.03
; SEQ ID NO 859
; LENGTH: 6558
; TYPE: DNA
; ORGANISM: Staphylococcus aureus
US-10-471-571A-859
```

```
Query Match      3.0%; Score 56.6; DB 6; Length 6558;
Best Local Similarity 45.7%; Pred. No. 0.0012;
Matches 197; Conservative 0; Mismatches 234; Indels 0; Gaps 0;

QY      545 ATATTATCTTAAATGTGTGATGCCAATGATGTGATTGATGTAAGTGAAGAAAGTAGAA 604
      5501 ATAAAGAAAGCGACAGATGAAGAAACAATAATGCAATAGCACAAAGTTGAAAAAGATTAA 5560
QY      605 CTGCTGAAAAATGAAGATTAAATGGCGAGATTTGTTATTCTTAAAGCGAAGCTCAC 664
      5561 TTAAAGCTTAAACAATAATGCTAGTGCAGTCAATGACAGATGGCATTTTAATTGC 5620
QY      665 CAGGACAGTAATAATATATCCAAAGATCAAAAGATGATTATATACATGACAAAGATA 724
      5621 ATGATGAGAAAAAAGCAATTTGCTGAATGAACTGTTATTAAAGAAAGCGTGGCTC 5680
QY      725 AAGAAACAAGCAAAAGCTTTATTACTGGGAAAAGTTATGAATTTGGCGAAAATGATTCA 784
      5681 GAGAACAATTGCAACATTTATTCACGATTAACAAAGCAATTAAGGAATATTCAAG 5740
QY      785 CAGAAAGAGAAATGAAACGGGAAATTACAGTTAAAAATCTTGAAGTTGGTGTGATA 844
      5741 CACCGTGAAGAAAGAAATATATATATGACAGTTACAAATATTTATGACACTGCTA 5800
QY      845 TTTTGAAGAAATAAAGCTCCAAATTAATGCAAAATTAATGCAAAATCAACAAACAC 904
      5801 TTGCAAAATGATCAAGTGTAGCAACAGTTGATTAACAGATCATTTAAATC 5860
QY      905 CATTTGCAATTGAAGCAAAACATCAAAACCTGTTGAAAAAAGCTCAAAATGATACCT 964
      5861 TACAAACAATACATGATTGATTTGATGTCATCTTATTAAGCCAGATGCTGAAAAAAGCA 5920
Db
```

```
QY      965 CTAAAGTTGAT 975
      5921 TTAATGATGAT 5931
Db
```

```
RESULT 3
US-11-218-305-3660
; Sequence 3660, Application US/11218305
; Publication No. US2006014195A1
; GENERAL INFORMATION:
; APPLICANT: MONSANTO TECHNOLOGY, LLC
; APPLICANT: McIsaid, Paul L.
; APPLICANT: Tao, Nenpheng
; APPLICANT: Wu, Kunsheng
; TITLE OF INVENTION: Polymorphic Markers and Methods of Genotyping
; FILE REFERENCE: Corn.
; CURRENT APPLICATION NUMBER: US/11/218,305
; PRIOR FILING DATE: 2005-09-01
; PRIOR APPLICATION NUMBER: US 60/606,880
; NUMBER OF SEQ ID NOS: 25043
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 3660
; LENGTH: 4009
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (22)..(22)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (111)..(111)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (178)..(178)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (285)..(285)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (290)..(290)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (308)..(308)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (310)..(313)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (341)..(341)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (362)..(364)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (408)..(408)
; OTHER INFORMATION: n is a, c, g, or t
US-11-218-305-3660
```

Query Match 2.9%; Score 54.6; DB 7; Length 4009;  
 Best Local Similarity 44.8%; Pred. No. 0.0031;  
 Matches 174; Conservative 0; Mismatches 214; Indels 0; Gaps 0;

Qy 649 AAAGCGAGGCTCCAGGCAAGTAAATATATATCAAGATCAAGATGATGATATAT 708  
 Db 3 AAAAAAAAAAGAGCGGNNAAAAAAAAAAAAAAAAAAAAAAAAAGAGAGAAAAA 62  
 Qy 709 ACATGACACCGATTAAGACAGCAAAACGCTTATATCTGGGAAAAATTATGAATT 768  
 Db 63 TAAAAAAGACAGCAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAA 122  
 Qy 769 GCGCAAAATGATTTCAAGAGCAAGATGAAAGCGGAAATTAAACATTTAAATCTT 828  
 Db 123 GGGGAAAGAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAA 182  
 Qy 829 GAGTTGTTCTATATTTTAAAGAGATGAAAGCTCAATATATGCAATTAATTGA 888  
 Db 183 AAAAAAAAAATATAATTTAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAA 242  
 Qy 889 AATCAACAAAAACACCATTTCAATTGAGCAAAACATCAACCTGTTAAAAACA 948  
 Db 243 AAAAAAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAA 302  
 Qy 949 GTCAAAATGATCTTAAGTTGATTAACACCAAGCTTATGATGTTAAAGATG 1008  
 Db 303 AAAAAAAAAAAAAAAAAAGAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAA 362  
 Qy 1009 GCAATTGGCGAAAAATTAATATCAAA 1036  
 Db 363 NNGAAGGGAAGAAAAAGAAAAAGAA 390

RESULT 4  
 US-11-218-305-24510/c  
 ; Sequence 24510, Application US/11218305  
 ; Publication No. US20060141495A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: MONSANTO TECHNOLOGY, LLC  
 ; APPLICANT: McLaidd, Paul L.  
 ; APPLICANT: Tao, Nengbing  
 ; APPLICANT: Wu, Kunsheng  
 ; TITLE OF INVENTION: Polymorphic Markers and Methods of Genotyping  
 ; TITLE OF INVENTION: Corn  
 ; FILE REFERENCE: 38-21 (53660)B  
 ; CURRENT APPLICATION NUMBER: US/11/218.305  
 ; PRIOR FILING DATE: 2005-09-01  
 ; PRIOR APPLICATION NUMBER: US 60/606,880  
 ; PRIOR FILING DATE: 2004-09-01  
 ; NUMBER OF SEQ ID NOS: 25043  
 ; SOFTWARE: PatentIn version 3.2  
 ; SEQ ID NO 24510  
 ; LENGTH: 2522  
 ; TYPE: DNA  
 ; ORGANISM: Zea mays  
 US-11-218-305-24510

Query Match 2.9%; Score 54.2; DB 7; Length 2522;  
 Best Local Similarity 45.1%; Pred. No. 0.0031;  
 Matches 200; Conservative 0; Mismatches 243; Indels 0; Gaps 0;

Qy 594 AAAGTAGAAGCTGTGAAAAATGAAGATTAATGCGCAATTTGTTATTTCTAAAG 653  
 Db 447 AAAAAATTAATATGCTTAACAACTTAATATGTAATATGTTATTAACGGAATGTAC 388  
 Qy 654 CGAAGCTCACGAGCAACAGTAAATATATCCAGAGAGTCAAGATGATTAATACATG 713  
 Db 387 CCGTTACAGTAGAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAA 328  
 Qy 714 GACAAAGATTAAGAAACAAAGCGCTTATTAAGTGGAAAAAGTTAATGAGCGCA 773  
 Db 327 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 268

Qy 774 AAATGATTTCAAGAGCAGAGATGAACGAGAAATTAACAGTTAAATCTTGAGGT 833  
 Db 267 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 208  
 Qy 834 TGGTTGTAATATTTTAAAGAAATGAAAGCTCAATATATGCAATTAATGAATCA 893  
 Db 207 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 148  
 Qy 894 AACAAAAACCATTTTCAATTAAGCAAAACATCAACCTGTTGAAAAACAGTCA 953  
 Db 147 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 88  
 Qy 954 AAATGATCTCTTAAGTTGATTAACACCAAGCTTATGATGTTAAAGATGCGCAT 1013  
 Db 87 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 28  
 Qy 1014 TGGCGAAAAATTAATATCAAA 1036  
 Db 27 AAAAAAAAAAAAAAAAAAAAAA 5

RESULT 5  
 US-11-218-305-12773/c  
 ; Sequence 12773, Application US/11218305  
 ; Publication No. US20060141495A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: MONSANTO TECHNOLOGY, LLC  
 ; APPLICANT: McLaidd, Paul L.  
 ; APPLICANT: Tao, Nengbing  
 ; APPLICANT: Wu, Kunsheng  
 ; TITLE OF INVENTION: Polymorphic Markers and Methods of Genotyping  
 ; TITLE OF INVENTION: Corn  
 ; FILE REFERENCE: 38-21 (53660)B  
 ; CURRENT APPLICATION NUMBER: US/11/218.305  
 ; PRIOR FILING DATE: 2005-09-01  
 ; PRIOR APPLICATION NUMBER: US 60/606,880  
 ; PRIOR FILING DATE: 2004-09-01  
 ; NUMBER OF SEQ ID NOS: 25043  
 ; SOFTWARE: PatentIn version 3.2  
 ; SEQ ID NO 12773  
 ; LENGTH: 2718  
 ; TYPE: DNA  
 ; ORGANISM: Zea mays

NAME/KEY: misc feature  
 LOCATION: (2410)..(2410)  
 OTHER INFORMATION: n is a, c, g, or t  
 FEATURE:  
 NAME/KEY: misc feature  
 LOCATION: (2425)..(2430)  
 OTHER INFORMATION: n is a, c, g, or t  
 FEATURE:  
 NAME/KEY: misc feature  
 LOCATION: (2434)..(2434)  
 OTHER INFORMATION: n is a, c, g, or t  
 FEATURE:  
 NAME/KEY: misc feature  
 LOCATION: (2442)..(2442)  
 OTHER INFORMATION: n is a, c, g, or t  
 FEATURE:  
 NAME/KEY: misc feature  
 LOCATION: (2504)..(2504)  
 OTHER INFORMATION: n is a, c, g, or t  
 FEATURE:  
 NAME/KEY: misc feature  
 LOCATION: (2510)..(2510)  
 OTHER INFORMATION: n is a, c, g, or t  
 FEATURE:  
 NAME/KEY: misc feature  
 LOCATION: (2570)..(2570)  
 OTHER INFORMATION: n is a, c, g, or t  
 FEATURE:

```
NAME/KEY: misc_feature
LOCATION: (2595)..(2599)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (2613)..(2613)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (2633)..(2633)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (2675)..(2675)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (2680)..(2680)
OTHER INFORMATION: n is a, c, g, or t
US-11-218-305-12773
```

Query Match 2.8%; Score 52; DB 7; Length 2718;

Best Local Similarity 44.3%; Pred. No. 0.011; Mismatches 228; Indels 0; Gaps 0;

Matches 181; Conservative 0; Mismatches 228; Indels 0; Gaps 0;

```
QY 585 ACATGTGAAAAAGTAGGAACTGCTGAAAAATGAGGATTAATGGCGCAGATTGTTAT 644
DB 2681 ANAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 2622
QY 645 TTCCTAAAGCGAGGCTCCAGGCAAGTAAATATATCCAGAGATCAAGATGAT 704
DB 2621 ATATATAAATATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 2562
QY 705 ATATCATGAGCAACGATTAAGCAAGCAAAAGCTTTATCTGGGAAAGTTATGA 764
DB 2561 GAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 2502
QY 765 AATTGCGCAAAATGATTTACAGAAAGAGATGAAACGGAGATTAACAGTTAAAA 824
DB 2501 AAAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 2442
QY 825 TCTTGAGGTGCTGCTATATTTTAAAGAGATTAAGCTCAAAATATGCAAGTTAT 884
DB 2441 GAGAGAGAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 2382
QY 885 TGAATAATGAAACAAACCACTTTCATTTGAAGCAATCAACACCTGTGAAA 944
DB 2381 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 2322
QY 945 AACAGTCAAAATGATACCTTAAGTTGATTAACCAACCAAGCTTA 993
DB 2321 GAAAAAGAGAAAGAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATA 2273
```

```
RESULT 6
US-11-216-545-774
Sequence 774, Application US/11216545
Publication No. US20060135758A1
GENERAL INFORMATION:
APPLICANT: MONSANTO Technology, LLC
APPLICANT: McElaird, Paul L
APPLICANT: Tao, Nengbing
APPLICANT: Wu, Kunsheng
TITLE OF INVENTION: Single Nucleotide Polymorphic Markers and Methods of Genotyping
FILE REFERENCE: 38-21 (33659)B
CURRENT APPLICATION NUMBER: US/11/216,545
PRIOR APPLICATION NUMBER: 2005-08-31
PRIOR FILING DATE: 2004-08-31
NUMBER OF SEQ ID NOS: 8783
SOFTWARE: PatentIn version 3.2
SEQ ID NO 774
```

```
LENGTH: 1826
TYPE: DNA
ORGANISM: Glycine max
FEATURE:
NAME/KEY: misc_feature
LOCATION: (10)..(10)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (15)..(16)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (31)..(31)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (33)..(39)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (42)..(43)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (45)..(46)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (51)..(53)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (58)..(61)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (73)..(73)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (75)..(80)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (97)..(97)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (100)..(103)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (106)..(109)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (112)..(113)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (118)..(121)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (125)..(125)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (133)..(133)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc_feature
```



LOCATION: (136)..(141)  
 OTHER INFORMATION: n is a, c, g, or t  
 FEATURE:  
 NAME/KEY: misc feature  
 LOCATION: (148)..(152)  
 OTHER INFORMATION: n is a, c, g, or t  
 FEATURE:  
 NAME/KEY: misc feature  
 LOCATION: (156)..(158)  
 OTHER INFORMATION: n is a, c, g, or t  
 FEATURE:  
 NAME/KEY: misc feature  
 LOCATION: (171)..(172)  
 OTHER INFORMATION: n is a, c, g, or t  
 FEATURE:  
 NAME/KEY: misc feature  
 LOCATION: (175)..(175)  
 OTHER INFORMATION: n is a, c, g, or t  
 FEATURE:  
 NAME/KEY: misc feature  
 LOCATION: (190)..(192)  
 OTHER INFORMATION: n is a, c, g, or t  
 FEATURE:  
 NAME/KEY: misc feature  
 LOCATION: (196)..(198)  
 OTHER INFORMATION: n is a, c, g, or t  
 FEATURE:  
 NAME/KEY: misc feature  
 LOCATION: (214)..(218)  
 OTHER INFORMATION: n is a, c, g, or t  
 FEATURE:  
 NAME/KEY: misc feature  
 LOCATION: (224)..(224)  
 OTHER INFORMATION: n is a, c, g, or t  
 FEATURE:  
 NAME/KEY: misc feature  
 LOCATION: (226)..(228)  
 OTHER INFORMATION: n is a, c, g, or t  
 FEATURE:  
 NAME/KEY: misc feature  
 LOCATION: (230)..(232)  
 OTHER INFORMATION: n is a, c, g, or t  
 FEATURE:  
 NAME/KEY: misc feature  
 LOCATION: (247)..(248)  
 OTHER INFORMATION: n is a, c, g, or t  
 FEATURE:  
 NAME/KEY: misc feature  
 LOCATION: (252)..(252)  
 OTHER INFORMATION: n is a, c, g, or t  
 FEATURE:  
 NAME/KEY: misc feature  
 LOCATION: (254)..(255)  
 OTHER INFORMATION: n is a, c, g, or t  
 FEATURE:  
 NAME/KEY: misc feature  
 LOCATION: (263)..(264)  
 OTHER INFORMATION: n is a, c, g, or t  
 FEATURE:  
 NAME/KEY: misc feature  
 LOCATION: (269)..(273)  
 OTHER INFORMATION: n is a, c, g, or t  
 FEATURE:  
 NAME/KEY: misc feature  
 LOCATION: (275)..(277)  
 OTHER INFORMATION: n is a, c, g, or t  
 FEATURE:  
 NAME/KEY: misc feature  
 LOCATION: (283)..(285)  
 OTHER INFORMATION: n is a, c, g, or t  
 FEATURE:  
 NAME/KEY: misc feature  
 LOCATION: (288)..(288)

OTHER INFORMATION: n is a, c, g, or t  
 FEATURE:  
 NAME/KEY: misc feature  
 LOCATION: (291)..(293)  
 OTHER INFORMATION: n is a, c, g, or t  
 FEATURE:  
 NAME/KEY: misc feature  
 LOCATION: (297)..(297)  
 OTHER INFORMATION: n is a, c, g, or t  
 FEATURE:  
 NAME/KEY: misc feature  
 LOCATION: (300)..(301)  
 OTHER INFORMATION: n is a, c, g, or t  
 FEATURE:  
 NAME/KEY: misc feature  
 LOCATION: (308)..(310)  
 OTHER INFORMATION: n is a, c, g, or t  
 FEATURE:  
 NAME/KEY: misc feature  
 LOCATION: (313)..(317)  
 OTHER INFORMATION: n is a, c, g, or t  
 FEATURE:  
 NAME/KEY: misc feature  
 LOCATION: (322)..(324)  
 OTHER INFORMATION: n is a, c, g, or t  
 FEATURE:  
 NAME/KEY: misc feature  
 LOCATION: (331)..(331)  
 OTHER INFORMATION: n is a, c, g, or t  
 FEATURE:  
 NAME/KEY: misc feature  
 LOCATION: (333)..(335)  
 OTHER INFORMATION: n is a, c, g, or t  
 FEATURE:  
 NAME/KEY: misc feature  
 LOCATION: (340)..(340)  
 OTHER INFORMATION: n is a, c, g, or t

Query Match 2.7%; Score 51.4; DB 8; Length 1826;  
 Best Local Similarity 29.4%; Pred. No. 0.013;  
 Matches 127; Conservative 0; Mismatches 305; Indels 0; Gaps 0;

QY 670 ACAGTAAATATATCCAGAGTCAAGATGATTTATATACATGACAAAGGATAAGAA 729  
 DB 20 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 79  
 QY 730 CAAGCAAAACGCTTATTACTGGAAAGTTATGAATTGGCGAAATGATTTACAGAA 789  
 DB 80 NAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 139  
 QY 790 GCAGAGATGGAACGGAGATTAACTGTAATTAAGTGGTTCGTATATTTA 849  
 DB 140 NAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 199  
 QY 850 GAAAGATGAAGCTCAATATATGCAATTAATGAAATCAACAAACACCATTT 909  
 DB 200 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 259  
 QY 910 ACAATTGAAGCAAAATCAACACCTGTTGAAAAAGTCAGCAAAATGATACCTTAA 969  
 DB 260 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 319  
 QY 970 GTTGATTAACACACCAAGCTTAGTGTGTAAGATGTGCAATTGGCGAAATTTAA 1029  
 DB 320 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 379  
 QY 1030 TATCAATTTCTGTAATATTCATGCGGATTCAGACAAAGGCGACGCTAATTA 1089  
 DB 380 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 439  
 QY 1090 TACGTCAAAATTC 1101  
 DB 440 AGAGTCAGCTAC 451

```
RESULT 7
US-10-517-441-691/c
; Sequence 691, Application US/10517441
; Publication No. US20060121467A1
; GENERAL INFORMATION:
; APPLICANT: POEKENS, John
; APPLICANT: HARBEC, Nadia
; APPLICANT: KOENIG, Thomas
; APPLICANT: MAIER, Sabine
; APPLICANT: MARTENS, John
; APPLICANT: MODEL, Fabian
; APPLICANT: NIMMICH, Inko
; APPLICANT: RUJAN, Tamas
; APPLICANT: SCHMITT, Armin
; APPLICANT: SCHMITT, Manfred
; APPLICANT: LOOK, Maxime P.
; APPLICANT: MARX, Almut
; APPLICANT: HOEFER, Heinz
; TITLE OF INVENTION: Method and nucleic acids for the improved treatment of breast cell
; TITLE OF INVENTION: Proliferative disorders
; FILE REFERENCE: 47675-93
; CURRENT APPLICATION NUMBER: US/10/517,441
; PRIOR FILING DATE: 2004-12-11
; PRIOR APPLICATION NUMBER: PCT/EP2003/010881
; PRIOR FILING DATE: 2003-10-01
; PRIOR APPLICATION NUMBER: DE 10317955.0
; PRIOR FILING DATE: 2003-04-17
; PRIOR APPLICATION NUMBER: DE 10300096.8
; PRIOR FILING DATE: 2003-01-07
; PRIOR APPLICATION NUMBER: DE 10245779.4
; PRIOR FILING DATE: 2002-10-01
; NUMBER OF SEQ ID NOS: 2147
; SEQ ID NO 691
; LENGTH: 5493
; TYPE: DNA
; ORGANISM: Artificial Sequence
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-517-441-691

Query Match      2.7%; Score 51.4; DB 6; Length 5493;
Best Local Similarity 46.1%; Pred. No. 0.019;
Matches 172; Conservative 0; Mismatches 201; Indels 0; Gaps 0;

QY 676 AATATATCCAGAGTCAAGATGATGATTATATACATGACACGAGTAAAGACAGCA 735
DB 5477 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAT 5418
QY 736 AAACCTTTATTTACTGGGAAAAGTTATGAAATTGGCGAAAATGATTTTCAGAGCAG 795
DB 5417 AAAAAAAAAACACAAAAAAATATCAACAAAAACAAACAAATATATATATATATAT 5358
QY 796 AATGACGCGAGATTAACAGTTAAATCTTGAGTTGTTCTGATTTTAAAGAA 855
DB 5357 AACAATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 5298
QY 856 GTAAAGCTCCAAATATATGAGATTAATTAATTAATTAATTAATTAATTAATTA 915
DB 5297 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAT 5238
QY 916 GAAGCAACATCAACACCTGTTGAAAAAAGTCACAAAAATGATTAATCTTAAGTTGAT 975
DB 5237 AAAAAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 5178
QY 976 AAAAAACACCAAGCTTGATGATTAAGATGCGAATGGGAAAAATTAATATCA 1035
DB 5177 CAAAAACAAAAAACAACAAACAAACAAACAAACAAACAAACAAACAAACAAAC 5118
QY 1036 ATTTCTGTAATA 1048
DB 5117 ACACCAAAAAA 5105
```

```
RESULT 8
US-10-471-571A-2917
; Sequence 2917, Application US/10471571A
; Publication No. US20060115490A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SPA
; TITLE OF INVENTION: STAPHYLOCOCCUS AUREUS PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE: P026927WO
; CURRENT APPLICATION NUMBER: US/10/471,571A
; PRIOR FILING DATE: 2003-09-12
; PRIOR APPLICATION NUMBER: GB-0107661.1
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 5642
; SOFTWARE: Seqwin99, version 1.03
; SEQ ID NO 2917
; LENGTH: 1407
; TYPE: DNA
; ORGANISM: Staphylococcus aureus
US-10-471-571A-2917

Query Match      2.7%; Score 50; DB 6; Length 1407;
Best Local Similarity 45.6%; Pred. No. 0.026;
Matches 176; Conservative 0; Mismatches 210; Indels 0; Gaps 0;

QY 754 AAAAGTTATGAATTTGGCGAAAATGATTTTCACAGACGAGATGAAACGGAGAAATTA 813
DB 322 AAAAAAGATTAACCCCTGAATTTGAAAGATTTATGAAAGAGGACATTTAAAGTCACCTTA 381
QY 814 ACAGTTAAAAATCTGAGGTTGTTGTTATTTTAAAGAAAGTAAAGCTCAATATAT 873
DB 382 GAATTAACAAATTTGAAATTAAGATTAATTTTAAAGTAAAGCTTTTATCAAACTAT 441
QY 874 GCAGATTAATTTGAAATTAACAAACAAACACCATTTTAAAGTAAAGCAACATCAACA 933
DB 442 AGAGATGATGTTGAAGTTATATAGTAAGTAAATTAATTAATTAATTAATTAATTAAT 501
QY 934 CCGTTGAAAAAAGCTGAAATTAATGATCTTAAGTTGATTAACAAACACCAAGCTTA 993
DB 502 GAAGAGCAAAATTAATTAAGCAAGTTTAAAGAAAGATTTGAAATTAATTAATTAATTA 561
QY 994 GATGTAAGATGTTGCGCAATTTGCGCAAAAAATTAATTAATTAATTTCTGTAATTAATTC 1053
DB 562 GAACCAATTAATTTGATTAATTTTATGATTAATTAATTAATTAATTAATTAATTTCT 621
QY 1054 TTGGGATTTGACAGCAAAAGAGCGCGCTAATTAATTAATTAATTAATTAATTTAGTAT 1113
DB 622 GTTTTAAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTT 681
QY 1114 AAACATGATGACGCTTAATCTTTGA 1139
DB 682 GACACTGAAGCAGCAAAAAAGTATGA 707

RESULT 9
US-10-517-441-413/c
; Sequence 413, Application US/10517441
; Publication No. US20060121467A1
; GENERAL INFORMATION:
; APPLICANT: POEKENS, John
; APPLICANT: HARBEC, Nadia
; APPLICANT: KOENIG, Thomas
; APPLICANT: MAIER, Sabine
; APPLICANT: MARTENS, John
; APPLICANT: MODEL, Fabian
; APPLICANT: NIMMICH, Inko
; APPLICANT: RUJAN, Tamas
; APPLICANT: SCHMITT, Armin
; APPLICANT: SCHMITT, Manfred
; APPLICANT: LOOK, Maxime P.
; APPLICANT: MARX, Almut
; APPLICANT: HOEFER, Heinz
```

```

; TITLE OF INVENTION: Method and nucleic acids for the improved treatment of breast cel
; FILE REFERENCE: 47675-93
; CURRENT APPLICATION NUMBER: US/10/517,441
; CURRENT FILING DATE: 2004-12-11
; PRIOR APPLICATION NUMBER: PCT/EP2003/010881
; PRIOR FILING DATE: 2003-10-01
; PRIOR APPLICATION NUMBER: DE 10317955.0
; PRIOR FILING DATE: 2003-04-17
; PRIOR APPLICATION NUMBER: DE 10300096.8
; PRIOR FILING DATE: 2003-01-07
; PRIOR APPLICATION NUMBER: DE 10245779.4
; PRIOR FILING DATE: 2002-10-01
; NUMBER OF SEQ ID NOS: 2147
; SEQ ID NO 413
; LENGTH: 19634
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-517-441-413

```

```

Query Match      2.6%; Score 49.8; DB 6; Length 19634;
Best Local Similarity 46.7%; Pred. No. 0.073;
Matches 192; Conservative 0; Mismatches 217; Indels 2; Gaps 1;

```

```

QY 705 ATATCATGACGACGATTAAGACCAAGCGCTTATTCTGGGAAAAGTTATGA 764
DB 6994 AAAAACTAATATAAAAAAATTAACAAAAAATTAACAAAAAATTAATGA 6935
QY 765 AATTGGCGAAATGATTTCAAGAGCAGAGATGGAACGGAGATTAACAGTTAAAA 824
DB 6934 AATT--CAATCATATCTATTAAAAATTAATTAATTAATTAATTAATTAATTA 6877
QY 825 TCTTAGGTTGGTGTATATTTTGAAGAGATAAGCTCCAAATATGACGAATTAAT 884
DB 6876 TAAAAACACTAAAAAATTAACATCTTACTAAAAAATTAATTAATTAATTAATTAAT 6817
QY 885 TGAATATCAACAAACCAACCATTTTCAATTTGAGCAACATCAACACCTGTTGAAA 944
DB 6816 CTATTAATAAAATATTAATCATTAATTAATTAATTAATTAATTAATTAATTAATTA 6757
QY 945 AACAGTCAAAAATGATACCTTAAGTTGATTAACCAACACCAAGCTTATGATGTAAGA 1004
DB 6756 AATATCAAAATTAATAAATTAACAAATTTTAATAAATTAACAAATTAATCAAAA 6697
QY 1005 TGTGCAATTGGCGAAAAATTAATTAATCAAAATTTCTGTAATTAATTCATTGGGATTC 1064
DB 6696 ATTTTAAATAATCTTAATTAATCAAAAATTTTAATAAATAATTAATAAATAATTA 6637
QY 1065 AGACAAAGAGCGACGCTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1115
DB 6636 ACAATATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 6586

```

RESULT 10  
US-10-517-441-687/c

```

; Sequence 687, Application US/10517441
; Publication No. US20060121467A1
; GENERAL INFORMATION:
; APPLICANT: FOCKENS, John
; APPLICANT: HARBEC, Nadia
; APPLICANT: KOENIG, Thomas
; APPLICANT: MAIER, Sabine
; APPLICANT: MARTENS, John
; APPLICANT: MODEL, Fabian
; APPLICANT: NIMMICH, Inko
; APPLICANT: RUJAN, Tamas
; APPLICANT: SCHMITT, Armin
; APPLICANT: SCHMITT, Manfred
; APPLICANT: LOOK, Maxime P.
; APPLICANT: MARX, Almutz
; APPLICANT: HOEFER, Heinz

```

```

; TITLE OF INVENTION: Method and nucleic acids for the improved treatment of breast cel
; FILE REFERENCE: 47675-93
; CURRENT APPLICATION NUMBER: US/10/517,441
; CURRENT FILING DATE: 2004-12-11
; PRIOR APPLICATION NUMBER: PCT/EP2003/010881
; PRIOR FILING DATE: 2003-10-01
; PRIOR APPLICATION NUMBER: DE 10317955.0
; PRIOR FILING DATE: 2003-04-17
; PRIOR APPLICATION NUMBER: DE 10300096.8
; PRIOR FILING DATE: 2003-01-07
; PRIOR APPLICATION NUMBER: DE 10245779.4
; PRIOR FILING DATE: 2002-10-01
; NUMBER OF SEQ ID NOS: 2147
; SEQ ID NO 687
; LENGTH: 19634
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-517-441-687

```

```

Query Match      2.6%; Score 49.8; DB 6; Length 19634;
Best Local Similarity 46.7%; Pred. No. 0.073;
Matches 192; Conservative 0; Mismatches 217; Indels 2; Gaps 1;

```

```

QY 705 ATATCATGACGACGATTAAGACCAAGCGCTTATTCTGGGAAAAGTTATGA 764
DB 6994 AAAAACTAATATAAAAAAATTAACAAAAAATTAACAAAAAATTAATGA 6935
QY 765 AATTGGCGAAATGATTTCAAGAGCAGAGATGGAACGGAGATTAACAGTTAAAA 824
DB 6934 AATT--CAATCATATCTATTAAAAATTAATTAATTAATTAATTAATTAATTA 6877
QY 825 TCTTAGGTTGGTGTATATTTTGAAGAGATAAGCTCCAAATATGACGAATTAAT 884
DB 6876 TAAAAACACTAAAAAATTAACATCTTACTAAAAAATTAATTAATTAATTAATTAAT 6817
QY 885 TGAATATCAACAAACCAACCATTTTCAATTTGAGCAACATCAACACCTGTTGAAA 944
DB 6816 CTATTAATAAAATATTAATCATTAATTAATTAATTAATTAATTAATTAATTAATTA 6757
QY 945 AACAGTCAAAAATGATACCTTAAGTTGATTAACCAACACCAAGCTTATGATGTAAGA 1004
DB 6756 AATATCAAAATTAATAAATTAACAAATTTTAATAAATTAACAAATTAATCAAAA 6697
QY 1005 TGTGCAATTGGCGAAAAATTAATTAATCAAAATTTCTGTAATTAATTCATTGGGATTC 1064
DB 6696 ATTTTAAATAATCTTAATTAATCAAAAATTTTAATAAATAATTAATAAATAATTA 6637
QY 1065 AGACAAAGAGCGACGCTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1115
DB 6636 ACAATATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 6586

```

RESULT 11  
US-11-266-748A-48160

```

; Sequence 48160, Application US/11266748A
; Publication No. US2006013463A1
; GENERAL INFORMATION:
; APPLICANT: Harkin, Paul
; APPLICANT: Johnston, Patrick
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcription Microarray Technology and
; FILE REFERENCE: 55815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266,748A
; CURRENT FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4

```

PRIOR FILING DATE: 2004-11-03  
PRIOR APPLICATION NUMBER: EP 04105507.0  
PRIOR FILING DATE: 2004-11-03  
PRIOR APPLICATION NUMBER: EP 04105485.9  
PRIOR FILING DATE: 2004-11-03  
PRIOR APPLICATION NUMBER: EP 04105484.2  
PRIOR FILING DATE: 2004-11-03  
PRIOR APPLICATION NUMBER: US 60/662,276  
PRIOR FILING DATE: 2005-03-14  
PRIOR APPLICATION NUMBER: US 60/700,293  
PRIOR FILING DATE: 2005-07-18  
NUMBER OF SEQ ID NOS: 483996  
SOFTWARE: PatentIn version 3.3  
SEQ ID NO 48160  
LENGTH: 951  
TYPE: DNA  
ORGANISM: Homo Sapiens  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (528)..(528)  
OTHER INFORMATION: n is a, c, g, or t  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (558)..(558)  
OTHER INFORMATION: n is a, c, g, or t  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (560)..(560)  
OTHER INFORMATION: n is a, c, g, or t  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (562)..(562)  
OTHER INFORMATION: n is a, c, g, or t  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (564)..(564)  
OTHER INFORMATION: n is a, c, g, or t  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (566)..(566)  
OTHER INFORMATION: n is a, c, g, or t  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (577)..(577)  
OTHER INFORMATION: n is a, c, g, or t  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (581)..(581)  
OTHER INFORMATION: n is a, c, g, or t  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (585)..(585)  
OTHER INFORMATION: n is a, c, g, or t  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (592)..(592)  
OTHER INFORMATION: n is a, c, g, or t  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (595)..(595)  
OTHER INFORMATION: n is a, c, g, or t  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (605)..(611)  
OTHER INFORMATION: n is a, c, g, or t  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (614)..(619)  
OTHER INFORMATION: n is a, c, g, or t  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (623)..(625)  
OTHER INFORMATION: n is a, c, g, or t

FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (630)..(631)  
OTHER INFORMATION: n is a, c, g, or t  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (635)..(636)  
OTHER INFORMATION: n is a, c, g, or t  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (638)..(639)  
OTHER INFORMATION: n is a, c, g, or t  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (641)..(641)  
OTHER INFORMATION: n is a, c, g, or t  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (656)..(661)  
OTHER INFORMATION: n is a, c, g, or t  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (666)..(675)  
OTHER INFORMATION: n is a, c, g, or t  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (685)..(685)  
OTHER INFORMATION: n is a, c, g, or t  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (687)..(698)  
OTHER INFORMATION: n is a, c, g, or t  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (701)..(701)  
OTHER INFORMATION: n is a, c, g, or t  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (705)..(705)  
OTHER INFORMATION: n is a, c, g, or t  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (707)..(707)  
OTHER INFORMATION: n is a, c, g, or t  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (712)..(712)  
OTHER INFORMATION: n is a, c, g, or t  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (715)..(716)  
OTHER INFORMATION: n is a, c, g, or t  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (725)..(725)  
OTHER INFORMATION: n is a, c, g, or t  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (732)..(733)  
OTHER INFORMATION: n is a, c, g, or t  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (738)..(739)  
OTHER INFORMATION: n is a, c, g, or t  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (741)..(743)  
OTHER INFORMATION: n is a, c, g, or t  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (746)..(746)  
OTHER INFORMATION: n is a, c, g, or t  
FEATURE:

```

; NAME/KEY: misc feature
; LOCATION: (761)..(762)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (768)..(769)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (771)..(771)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (773)..(773)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (782)..(782)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (800)..(800)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (802)..(805)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (821)..(824)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (826)..(831)

```

Query Match 2.6%; Score 49.6; DB 8; Length 951;  
 Best Local Similarity 30.8%; Pred. No. 0.028;  
 Matches 121; Conservative 0; Mismatches 272; Indels 0; Gaps 0;

```

Qy 612 AATGAGATTAATGCGGCGAATTTCTTAATTAAGCGAGCTCAGCGAC 671
Db 559 AATGAGATTAATGCGGCGAATTTCTTAATTAAGCGAGCTCAGCGAC 618
Qy 672 AGTAATATATATCCAGGAGTCAAGATGATTATATACAGCAACGATTAAGACA 731
Db 619 AATGAGATTAATGCGGCGAATTTCTTAATTAAGCGAGCTCAGCGAC 678
Qy 732 AGCAAAAGCTTATTAATCTGCGAAAGTTATGAAATTCGCGAAATGATTTCAAGAAC 791
Db 679 AATGAGATTAATGCGGCGAATTTCTTAATTAAGCGAGCTCAGCGAC 738
Qy 792 AGAGATGAGGCGGAGATTAAACGTTAAATCTGAGGTTGCTGATTAATTTAGA 851
Db 739 AATGAGATTAATGCGGCGAATTTCTTAATTAAGCGAGCTCAGCGAC 798
Qy 852 AGAGATTAATGCGGAGATTAAACGTTAAATCTGAGGTTGCTGATTAATTTAGA 911
Db 799 AATGAGATTAATGCGGAGATTAAACGTTAAATCTGAGGTTGCTGATTAATTTAGA 858
Qy 912 AATGAGATTAATGCGGAGATTAAACGTTAAATCTGAGGTTGCTGATTAATTTAGA 971
Db 859 AATGAGATTAATGCGGAGATTAAACGTTAAATCTGAGGTTGCTGATTAATTTAGA 918
Qy 972 TGATTAATCAACACGAGCTTAGATGTTAAGA 1004
Db 919 AATGAGATTAATGCGGAGATTAAACGTTAAATCTGAGGTTGCTGATTAATTTAGA 951

```

RESULT 12

```

US-10-517-441-417/c
; Sequence 417, Application US/10517441
; Publication No. US20060121467A1
; GENERAL INFORMATION:
; APPLICANT: FOKKENS, John
; APPLICANT: HARBESCK, Nadia
; APPLICANT: KOENIG, Thomas
; APPLICANT: MAIER, Sabine
; APPLICANT: MARTENS, John
; APPLICANT: MODELL, Fabian
; APPLICANT: NIMMERICH, Inko
; APPLICANT: RUJAN, Tamas
; APPLICANT: SCHMITT, Armin
; APPLICANT: SCHMITT, Manfred
; APPLICANT: LOOK, Maxime F.
; APPLICANT: MARK, Almuth
; APPLICANT: HOEPLER, Heinz
; TITLE OF INVENTION: Method and nucleic acids for the improved treatment of breast cell
; FILE REFERENCE: 47675-93
; CURRENT APPLICATION NUMBER: US/10/517,441
; PRIOR FILING DATE: 2004-12-11
; PRIOR APPLICATION NUMBER: PCT/EP2003/010881
; PRIOR FILING DATE: 2003-10-01
; PRIOR APPLICATION NUMBER: DE 10317955.0
; PRIOR FILING DATE: 2003-04-17
; PRIOR APPLICATION NUMBER: DE 10300096.8
; PRIOR FILING DATE: 2003-01-07
; PRIOR APPLICATION NUMBER: DE 10245779.4
; PRIOR FILING DATE: 2002-10-01
; NUMBER OF SEQ ID NOS: 2147
; SEQ ID NO 417
; LENGTH: 5493
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-517-441-417

```

Query Match 2.6%; Score 49.6; DB 6; Length 5493;  
 Best Local Similarity 44.3%; Pred. No. 0.052;  
 Matches 202; Conservative 0; Mismatches 254; Indels 0; Gaps 0;

```

Qy 648 TAAAGCGAGCTCAGGAGCACTGTAATTAATTCAGAGATCAAGATGATTA 707
Db 5370 TAAAGCGAGCTCAGGAGCACTGTAATTAATTCAGAGATCAAGATGATTA 5311
Qy 708 TACATGAGCAACGAGTAAGAACAGCAACGCTTTATTAATGAGTAAT 767
Db 5310 TAAAGCGAGCTCAGGAGCACTGTAATTAATTCAGAGATCAAGATGATTA 5251
Qy 768 TGGCGAATATGATTTACAGAGAGAGATGAAACGAGATTAACGTTAAATCT 827
Db 5250 TAAAGCGAGCTCAGGAGCACTGTAATTAATTCAGAGATCAAGATGATTA 5191
Qy 828 TGAAGTGGTGTATATTTAGAGAGTAATTAACCTCAATTAATGCAATTAATTA 887
Db 5190 TAAAGCGAGCTCAGGAGCACTGTAATTAATTCAGAGATCAAGATGATTA 5131
Qy 888 AATCAACCAAAACACCAATTTACATTTGAAGCAACATCAACACCTGTTGAAGAAC 947
Db 5130 AATCAACCAAAACACCAATTTACATTTGAAGCAACATCAACACCTGTTGAAGAAC 5071
Qy 948 AGTCAGCAATGATCTGTAATGATTAATTAATTAATTAATTAATTAATTAATTA 1007
Db 5070 AATCAACCAAAACACCAATTTACATTTGAAGCAACATCAACACCTGTTGAAGAAC 5011
Qy 1008 GGCATTTGGCGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1067
Db 5010 AATCAACCAAAACACCAATTTACATTTGAAGCAACATCAACACCTGTTGAAGAAC 4951
Qy 1068 CAAAGAGCGAGCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1103

```

Db 4950 AATAAAACACCACAAATAATCAATAATATAA 4915

RESULT 13

US-11-266-748A-222640

Sequence 222640, Application US/11266748A

Publication No. US20060134663A1

GENERAL INFORMATION:

APPLICANT: Harkin, Paul

APPLICANT: Mulligan, Karl

TITLE OF INVENTION: Transcriptome Microarray Technology and

TITLE OF INVENTION: Methods of Using the Same

FILE REFERENCE: 55815-0102 (319189)

CURRENT APPLICATION NUMBER: US/11/266,748A

CURRENT FILING DATE: 2005-11-03

PRIOR APPLICATION NUMBER: EP 04105479.2

PRIOR FILING DATE: 2004-11-03

PRIOR APPLICATION NUMBER: EP 04105482.6

PRIOR FILING DATE: 2004-11-03

PRIOR APPLICATION NUMBER: EP 04105483.4

PRIOR FILING DATE: 2004-11-03

PRIOR APPLICATION NUMBER: EP 04105507.0

PRIOR FILING DATE: 2004-11-03

PRIOR APPLICATION NUMBER: EP 04105485.9

PRIOR FILING DATE: 2004-11-03

PRIOR APPLICATION NUMBER: EP 04105484.2

PRIOR FILING DATE: 2004-11-03

PRIOR APPLICATION NUMBER: US 60/662,276

PRIOR FILING DATE: 2005-03-14

PRIOR APPLICATION NUMBER: US 60/700,293

PRIOR FILING DATE: 2005-07-18

NUMBER OF SEQ ID NOS: 483996

SOFTWARE: PatentIn version 3.3

SEQ ID NO 222640

LENGTH: 1000

TYPE: DNA

ORGANISM: Homo Sapiens

US-11-266-748A-222640

Query Match 2.6%; Score 49.2; DB 8; Length 1000;

Best Local Similarity 50.9%; Pred. No. 0.036;

Matches 117; Conservative 0; Mismatches 113; Indels 0; Gaps 0;

QY 807 AGAATTAACAGTTAAATCTTGAGTGTGCTGTAATATTAGAGAAAGTAAAGCTCC 866

DB 757 ACATTAACTGTGTAATGTAATGTAAGAAAGCTCAGCATGTGTCATAAATACTGTCC 816

QY 867 AAATTAATGAGAAATTAATGAAATCAAAACAAACACCATTTACATTTGAAGCAACA 926

DB 817 TAATTCATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 876

QY 927 TCAACACCTGTGTAAGAAAGCTCAAAATGATACCTTAAGTTGATTAACACACACC 986

DB 877 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 936

QY 987 AAGCTTAAGTGAAGATGTCGCAATTTGGCGAAATAATTAATATCAAA 1036

DB 937 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 986

RESULT 14

US-11-266-748A-289165

Sequence 289165, Application US/11266748A

Publication No. US20060134663A1

GENERAL INFORMATION:

APPLICANT: Harkin, Paul

APPLICANT: Johnston, Patrick

APPLICANT: Mulligan, Karl

TITLE OF INVENTION: Transcriptome Microarray Technology and

TITLE OF INVENTION: Methods of Using the Same

FILE REFERENCE: 55815-0102 (319189)

CURRENT APPLICATION NUMBER: US/11/266,748A

CURRENT FILING DATE: 2005-11-03

PRIOR APPLICATION NUMBER: EP 04105479.2

PRIOR FILING DATE: 2004-11-03

PRIOR APPLICATION NUMBER: EP 04105482.6

PRIOR FILING DATE: 2004-11-03

PRIOR APPLICATION NUMBER: EP 04105483.4

PRIOR FILING DATE: 2004-11-03

PRIOR APPLICATION NUMBER: EP 04105507.0

PRIOR FILING DATE: 2004-11-03

PRIOR APPLICATION NUMBER: EP 04105485.9

PRIOR FILING DATE: 2004-11-03

PRIOR APPLICATION NUMBER: EP 04105484.2

PRIOR FILING DATE: 2004-11-03

PRIOR APPLICATION NUMBER: US 60/662,276

PRIOR FILING DATE: 2005-03-14

PRIOR APPLICATION NUMBER: US 60/700,293

PRIOR FILING DATE: 2005-07-18

NUMBER OF SEQ ID NOS: 483996

SOFTWARE: PatentIn version 3.3

SEQ ID NO 289165

LENGTH: 1000

TYPE: DNA

ORGANISM: Homo Sapiens

US-11-266-748A-289165

Query Match 2.6%; Score 49.2; DB 8; Length 1000;

Best Local Similarity 50.9%; Pred. No. 0.036;

Matches 117; Conservative 0; Mismatches 113; Indels 0; Gaps 0;

QY 807 AGAATTAACAGTTAAATCTTGAGTGTGCTGTAATATTAGAGAAAGTAAAGCTCC 866

DB 757 ACATTAACTGTGTAATGTAATGTAAGAAAGCTCAGCATGTGTCATAAATACTGTCC 816

QY 867 AAATTAATGAGAAATTAATGAAATCAAAACAAACACCATTTCAATTTGAAGCAACA 926

DB 817 TAATTCATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 876

QY 927 TCAACACCTGTGTAAGAAAGCTCAAAATGATACCTTAAGTTGATTAACACACACC 986

DB 877 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 936

QY 987 AAGCTTAAGTGAAGATGTCGCAATTTGGCGAAATAATTAATATCAAA 1036

DB 937 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 986

RESULT 15

US-11-266-748A-340594/C

Sequence 340594, Application US/11266748A

Publication No. US20060134663A1

GENERAL INFORMATION:

APPLICANT: Harkin, Paul

APPLICANT: Johnston, Patrick

APPLICANT: Mulligan, Karl

TITLE OF INVENTION: Transcriptome Microarray Technology and

TITLE OF INVENTION: Methods of Using the Same

FILE REFERENCE: 55815-0102 (319189)

CURRENT APPLICATION NUMBER: US/11/266,748A

CURRENT FILING DATE: 2005-11-03

PRIOR APPLICATION NUMBER: EP 04105479.2

PRIOR FILING DATE: 2004-11-03

PRIOR APPLICATION NUMBER: EP 04105482.6

PRIOR FILING DATE: 2004-11-03

PRIOR APPLICATION NUMBER: EP 04105483.4

PRIOR FILING DATE: 2004-11-03

PRIOR APPLICATION NUMBER: EP 04105507.0

PRIOR FILING DATE: 2004-11-03

PRIOR APPLICATION NUMBER: EP 04105485.9

PRIOR FILING DATE: 2004-11-03

PRIOR APPLICATION NUMBER: EP 04105484.2

PRIOR FILING DATE: 2004-11-03

PRIOR APPLICATION NUMBER: US 60/662,276

PRIOR FILING DATE: 2005-03-14

; PRIOR APPLICATION NUMBER: US 60/700,293  
 ; PRIOR FILING DATE: 2005-07-18  
 ; NUMBER OF SEQ ID NOS: 483996  
 ; SOFTWARE: Patentin version 3.3  
 ; SEQ ID NO 340594  
 ; LENGTH: 1000  
 ; TYPE: DNA  
 ; ORGANISM: Homo Sapiens  
 US-11-266-748A-340594

Query Match 2.6%; Score 49.2; DB 8; Length 1000;  
 Best Local Similarity 50.9%; Pred. No. 0.036;  
 Matches 117; Conservative 0; Mismatches 113; Indels 0; Gaps 0;

|    |     |                                                           |      |
|----|-----|-----------------------------------------------------------|------|
| QY | 807 | AGATTACAGTTAAATCTTGAGGTGTCGTATATTGAAAGAGTAAAGCTCC         | 866  |
| DB | 244 | ACATTAACTGTGTAAATGTAAGAGAGCTCAGCATGTGTCAACATTAAGCTGTC     | 185  |
| QY | 867 | AAATATGAGAAATTAATGAAATCAAAACAAACACCATTTACATTTGAGCAACAA    | 926  |
| DB | 184 | TAAATCTCAAAATTAATGAAATCAAAACAAACACCATTTACATTTGAGCAACAA    | 125  |
| QY | 927 | TCAAAACACCTGTGAAAAACAGTCAAAATGATACCTCTAAAGTTGATTAACACACC  | 986  |
| DB | 124 | AAAAATTAATGAAATCAAAACAGTCAAAATGATACCTCTAAAGTTGATTAACACACC | 65   |
| QY | 987 | AAGCTTAAGTGTAAAGATGTGCAATTGGCAAAAAATTAAATATCAAA           | 1036 |
| DB | 64  | AAAAATTAATGAAATCAAAACAGTCAAAATGATACCTCTAAAGTTGATTAACACACC | 15   |

Search completed: July 6, 2006, 00:51:23  
 Job time : 177 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.9  
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 5, 2006, 20:23:58 ; Search time 265 Seconds

(without alignments)  
13281.342 Million cell updates/sec

Title: US-10-661-809a-12

Perfect score: 1881  
Sequence: 1 atgaagcaatttaaaagct.....gacgtgagaagaagaatgct 1881

Scoring table: IDENTITY\_NUC

Gapop 10.0, Gapext 1.0

Searched: 1403666 seqs, 93554401 residues

Total number of hits satisfying chosen parameters: 2807332

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Issued Patents NA:\*

1: /EMC Celerra\_SIDS3/ptodata/2/ina/1 COMB.seq:\*\n2: /EMC Celerra\_SIDS3/ptodata/2/ina/5 COMB.seq:\*\n3: /EMC Celerra\_SIDS3/ptodata/2/ina/6 COMB.seq:\*\n4: /EMC Celerra\_SIDS3/ptodata/2/ina/6 COMB.seq:\*\n5: /EMC Celerra\_SIDS3/ptodata/2/ina/7 COMB.seq:\*\n6: /EMC Celerra\_SIDS3/ptodata/2/ina/7 COMB.seq:\*\n7: /EMC Celerra\_SIDS3/ptodata/2/ina/PTUS.COMB.seq:\*\n8: /EMC Celerra\_SIDS3/ptodata/2/ina/PTUS.COMB.seq:\*\n9: /EMC Celerra\_SIDS3/ptodata/2/ina/RE.COMB.seq:\*\n10: /EMC Celerra\_SIDS3/ptodata/2/ina/backfile1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score  | Query Match | Length  | DB ID | Description          |
|------------|--------|-------------|---------|-------|----------------------|
| 1          | 1881   | 100.0       | 1980    | 3     | US-09-134-000C-2719  |
| 2          | 1881   | 100.0       | 1983    | 3     | US-09-071-035-217    |
| 3          | 1881   | 100.0       | 1983    | 3     | US-10-206-576-217    |
| 4          | 1881   | 100.0       | 1983    | 3     | US-09-071-035-219    |
| 5          | 1687   | 89.7        | 1687    | 3     | US-10-206-576-219    |
| 6          | 1014.6 | 53.9        | 1995    | 3     | US-09-107-532A-3598  |
| 7          | 74.6   | 4.0         | 7218    | 2     | US-08-232-463-14     |
| 8          | 64.6   | 3.4         | 6609    | 3     | US-10-172-502-1      |
| 9          | 57.2   | 3.0         | 1039    | 3     | US-09-902-540-1880   |
| 10         | 57     | 3.0         | 612     | 3     | US-08-956-171E-171   |
| 11         | 56.6   | 3.0         | 11126   | 3     | US-08-956-171E-171   |
| 12         | 56.6   | 3.0         | 11126   | 3     | US-08-956-171E-171   |
| 13         | 56.4   | 3.0         | 1664976 | 3     | US-08-916-421B-1     |
| 14         | 56.4   | 3.0         | 1664976 | 3     | US-08-916-421B-1     |
| 15         | 54     | 2.9         | 1141    | 3     | US-09-692-570-1      |
| 16         | 53.8   | 2.9         | 1434    | 3     | US-09-806-708B-22    |
| 17         | 53     | 2.8         | 3942    | 3     | US-09-134-000C-2718  |
| 18         | 52.2   | 2.8         | 5361    | 3     | US-08-973-462-2      |
| 19         | 52.2   | 2.8         | 6152    | 3     | US-08-973-462-1      |
| 20         | 52.2   | 2.8         | 18016   | 3     | US-09-949-016-15487  |
| 21         | 52     | 2.8         | 1898    | 3     | US-09-769-787-339    |
| 22         | 52     | 2.8         | 20986   | 3     | US-08-961-527-54     |
| 23         | 51.8   | 2.8         | 601     | 3     | US-09-949-016-133019 |

|    |      |     |         |   |                     |                    |
|----|------|-----|---------|---|---------------------|--------------------|
| 24 | 51.8 | 2.8 | 1891    | 3 | US-08-973-462-3     | Sequence 3, Appl1  |
| 25 | 51.6 | 2.7 | 614     | 3 | US-08-902-540-1318  | Sequence 1316, Ap  |
| 26 | 51.6 | 2.7 | 5181    | 2 | US-08-257-073-10    | Sequence 10, Appl1 |
| 27 | 51   | 2.7 | 2223    | 2 | US-08-257-073-4     | Sequence 4, Appl1  |
| 28 | 51   | 2.7 | 640681  | 3 | US-09-790-988-1     | Sequence 1, Appl1  |
| 29 | 50.8 | 2.7 | 963     | 3 | US-09-134-000C-2132 | Sequence 2132, Ap  |
| 30 | 50.6 | 2.7 | 1305    | 3 | US-09-329-234A-6    | Sequence 6, Appl1  |
| 31 | 50.4 | 2.7 | 1664976 | 3 | US-08-916-421B-1    | Sequence 1, Appl1  |
| 32 | 50.4 | 2.7 | 1664976 | 3 | US-08-916-421B-1    | Sequence 1, Appl1  |
| 33 | 50   | 2.7 | 1353    | 3 | US-08-956-171E-521  | Sequence 521, App  |
| 34 | 50   | 2.7 | 1353    | 3 | US-08-781-986A-521  | Sequence 521, App  |
| 35 | 49.8 | 2.6 | 699     | 3 | US-09-461-697-193   | Sequence 193, App  |
| 36 | 49.8 | 2.6 | 699     | 3 | US-09-461-697-193   | Sequence 193, App  |
| 37 | 49.8 | 2.6 | 717     | 3 | US-09-461-697-187   | Sequence 187, App  |
| 38 | 49.8 | 2.6 | 717     | 3 | US-09-461-697-187   | Sequence 187, App  |
| 39 | 49.8 | 2.6 | 819     | 3 | US-09-461-697-185   | Sequence 185, App  |
| 40 | 49.8 | 2.6 | 1669    | 3 | US-09-461-697-184   | Sequence 184, App  |
| 41 | 49.8 | 2.6 | 1983    | 3 | US-09-107-532A-3061 | Sequence 3061, Ap  |
| 42 | 49.8 | 2.6 | 50000   | 3 | US-09-662-254B-25   | Sequence 25, Appl1 |
| 43 | 49.6 | 2.6 | 5887    | 4 | US-09-747-385-6     | Sequence 6, Appl1  |
| 44 | 49.6 | 2.6 | 9955    | 4 | US-09-747-385-15    | Sequence 15, Appl1 |
| 45 | 49.6 | 2.6 | 1116    | 3 | US-09-248-796A-9660 | Sequence 9660, Ap  |

## ALIGNMENTS

|                                                                         |     |              |                                                |     |  |  |  |  |  |
|-------------------------------------------------------------------------|-----|--------------|------------------------------------------------|-----|--|--|--|--|--|
| RESULT 1                                                                |     |              |                                                |     |  |  |  |  |  |
| US-09-134-000C-2719                                                     |     |              |                                                |     |  |  |  |  |  |
| ; Sequence 2719, Application US/09134000C                               |     |              |                                                |     |  |  |  |  |  |
| ; Patent No. 6617156                                                    |     |              |                                                |     |  |  |  |  |  |
| ; GENERAL INFORMATION:                                                  |     |              |                                                |     |  |  |  |  |  |
| ; APPLICANT: Lynn Doucette-Stamm et al                                  |     |              |                                                |     |  |  |  |  |  |
| ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO |     |              |                                                |     |  |  |  |  |  |
| ; FILE REFERENCE: 032796-032                                            |     |              |                                                |     |  |  |  |  |  |
| ; CURRENT APPLICATION NUMBER: US/09/134,000C                            |     |              |                                                |     |  |  |  |  |  |
| ; PRIOR FILING DATE: 1998-08-13                                         |     |              |                                                |     |  |  |  |  |  |
| ; PRIOR APPLICATION NUMBER: US 60/055,778                               |     |              |                                                |     |  |  |  |  |  |
| ; PRIOR FILING DATE: 1997-08-15                                         |     |              |                                                |     |  |  |  |  |  |
| ; NUMBER OF SEQ ID NOS: 6812                                            |     |              |                                                |     |  |  |  |  |  |
| ; SOFTWARE: PatentIn version 3.1                                        |     |              |                                                |     |  |  |  |  |  |
| ; SEQ ID NO 2719                                                        |     |              |                                                |     |  |  |  |  |  |
| ; LENGTH: 1980                                                          |     |              |                                                |     |  |  |  |  |  |
| ; TYPE: DNA                                                             |     |              |                                                |     |  |  |  |  |  |
| ; ORGANISM: Enterococcus faecalis                                       |     |              |                                                |     |  |  |  |  |  |
| US-09-134-000C-2719                                                     |     |              |                                                |     |  |  |  |  |  |
| Query Match                                                             |     |              |                                                |     |  |  |  |  |  |
| Best Local Similarity 100.0%; Score 1881; DB 3; Length 1980;            |     |              |                                                |     |  |  |  |  |  |
| Matches 1881; Conservative 0; Mismatches 0; Indels 0; Gaps 0;           |     |              |                                                |     |  |  |  |  |  |
| QY                                                                      | 1   | ATGAAGCAATTA | AAAAAGTTGTGACACCGTAGTACCTGTGTAATTTGCCACTT      | 60  |  |  |  |  |  |
| DB                                                                      | 97  | ATGAAGCAATTA | AAAAAGTTGTGACACCGTAGTACCTGTGTAATTTGCCACTT      | 156 |  |  |  |  |  |
| QY                                                                      | 61  | TTCACAACTGTT | TGGGACCAACCTGATTTGGAGAAAGGAGGAGCGGACAG         | 120 |  |  |  |  |  |
| DB                                                                      | 157 | TTCACAACTGTT | TGGGACCAACCTGATTTGGAGAAAGGAGGAGCGGACAG         | 216 |  |  |  |  |  |
| QY                                                                      | 121 | CTCGGATTCACA | AAAAAGTAATGACGATTTACAGATCCGCTTAATTCAAATAGCGGG  | 180 |  |  |  |  |  |
| DB                                                                      | 217 | CTCGGATTCACA | AAAAAGTAATGACGATTTACAGATCCGCTTAATTCAAATAGCGGG  | 276 |  |  |  |  |  |
| QY                                                                      | 181 | AAAGAAATGAG  | CGATTTGATTAATATGACAGACTGGGAGATGTGACGTTAGTATTAT | 240 |  |  |  |  |  |
| DB                                                                      | 277 | AAAGAAATGAG  | CGATTTGATTAATATGACAGACTGGGAGATGTGACGTTAGTATTAT | 336 |  |  |  |  |  |
| QY                                                                      | 241 | AAACGAGGAAC  | GAATTTTACAGAGACGAGCGGACGAGGCTTGAATGACACTAA     | 300 |  |  |  |  |  |
| DB                                                                      | 337 | AAACGAGGAAC  | GAATTTTACAGAGACGAGCGGACGAGGCTTGAATGACACTAA     | 396 |  |  |  |  |  |
| QY                                                                      | 301 | CAAGCTGTCCA  | AAAGTTTAATCTCTGGAAACCTGTGCTCAAGGAACACGATCAAT   | 360 |  |  |  |  |  |

Db 397 CAAGCTGTCGAAAGTTTAACTCTCGGAAACCTGTGCTCAGAGAACCCAGTCGAANT 456  
Qy 361 GGGAAATGTCAGTTCAGTTACCTTAAACAAAATGGTAAAGATGCACTGTATACATT 420  
Db 457 GGGAAATGTCAGTTCAGTTACCTTAAACAAAATGGTAAAGATGCACTGTATACATT 516  
Qy 421 AAAGAAACCAAAAGAGGTGTAGTGTCTACGAAATATNGTGTGCGCTTCCAGTT 480  
Db 517 AAAGAAACCAAAAGAGGTGTAGTGTCTACGAAATATNGTGTGCGCTTCCAGTT 576  
Qy 481 TACGAATGATCAAGCAAAACAGATGTTCTTAAATATGAAACAGAAATTAAGCGTT 540  
Db 577 TACGAATGATCAAGCAAAACAGATGTTCTTAAATATGAAACAGAAATTAAGCGTT 636  
Qy 541 GTTCATATTTATCTTAAAAATGTGTAGCGCAATGATGTAGTTTACATGTGAAAAAGTA 600  
Db 637 GTTCATATTTATCTTAAAAATGTGTAGCGCAATGATGTAGTTTACATGTGAAAAAGTA 696  
Qy 601 GGAATGCTGAAATGAAAGATTTAAATGGCGGAGAAATTTGTTATTTCTTAAAGCGAAGGC 660  
Db 697 GGAATGCTGAAATGAAAGATTTAAATGGCGGAGAAATTTGTTATTTCTTAAAGCGAAGGC 756  
Qy 661 TCACCAAGGCAAGTAAATATATCCAGAGATCAAAAGATGATTAATATCATGACAAAG 720  
Db 757 TCACCAAGGCAAGTAAATATATCCAGAGATCAAAAGATGATTAATATCATGACAAAG 816  
Qy 721 GATTAAGAACAGCAAAACGCTTTATTAATCTGCGAAATGTTATGAAATGGCGAATATGAT 780  
Db 817 GATTAAGAACAGCAAAACGCTTTATTAATCTGCGAAATGTTATGAAATGGCGAATATGAT 876  
Qy 781 TTCACAGAGCAGAGATGGAACGGGAGAAATTAACGTTAAAAATCTTGAGGTGGTTGG 840  
Db 877 TTCACAGAGCAGAGATGGAACGGGAGAAATTAACGTTAAAAATCTTGAGGTGGTTGG 936  
Qy 841 TATATTTTGAAGAGTAAAGCTCCAAATTAATGCAATTAATGAAATCAAAACAAA 900  
Db 937 TATATTTTGAAGAGTAAAGCTCCAAATTAATGCAATTAATGAAATCAAAACAAA 996  
Qy 901 ACACCATTTTCAATTTGAGCAAAACATCAAAACCTGTGTAAGAAAAACAGTCAAAATGAT 960  
Db 997 ACACCATTTTCAATTTGAGCAAAACATCAAAACCTGTGTAAGAAAAACAGTCAAAATGAT 1056  
Qy 961 ACCTCTAAAGTTGATTAACCAACCAACCTTATGATGTTAAAGATGTGGCAATTTGGCGAA 1020  
Db 1057 ACCTCTAAAGTTGATTAACCAACCAACCTTATGATGTTAAAGATGTGGCAATTTGGCGAA 1116  
Qy 1021 AAAATTTAATATCAAAATTTCTGTAATATTTCAATTTGGGATTTGCAGACAAAGGCGAC 1080  
Db 1117 AAAATTTAATATCAAAATTTCTGTAATATTTCAATTTGGGATTTGCAGACAAAGGCGAC 1176  
Qy 1081 GCTAATTAATACGTCAAAATTTCAATTTAGTTGTAATAACATGACAGCTTTAATCTTTGAT 1140  
Db 1177 GCTAATTAATACGTCAAAATTTCAATTTAGTTGTAATAACATGACAGCTTTAATCTTTGAT 1236  
Qy 1141 AACGTGACTTCTGAGAGATGATGTTATGCGTTATATGATGGGATACAGTATGCTCTCT 1200  
Db 1237 AACGTGACTTCTGAGAGATGATGTTATGCGTTATATGATGGGATACAGTATGCTCTCT 1296  
Qy 1201 GAAATTTATCAAGTGACTGAACCAAGCAAAATGGCTTCACTGTGCGCTTATCAAGCGTAT 1260  
Db 1297 GAAATTTATCAAGTGACTGAACCAAGCAAAATGGCTTCACTGTGCGCTTATCAAGCGTAT 1356  
Qy 1261 ATTCTTACGCTTAACACAGGCGGACACTTAAATTTGTTTACTTTATGCAATTTAAATGAA 1320  
Db 1357 ATTCTTACGCTTAACACAGGCGGACACTTAAATTTGTTTACTTTATGCAATTTAAATGAA 1416  
Qy 1321 AAAGCAGATCTTACGAAAGGCTTTAAATTAAGAGCGCAATGTTGATAACGCTATACCGAC 1380  
Db 1417 AAAGCAGATCTTACGAAAGGCTTTAAATTAAGAGCGCAATGTTGATAACGCTATACCGAC 1476  
Qy 1381 GACCAAAACCAACCACTGTGTAAGTTGTGACAGGTGGGAAACGTTTCAATTAAGTCGAT 1440

Db 1477 GACCAAAACCAACCACTGTGTAAGTTGTGACAGGTGGGAAACGTTTCAATTAAGTCGAT 1536  
Qy 1441 GGCATGTGACAGCGACACCAAGCCTTGGCGGAGCTTCTTTGTGTCCGTGATCAAAAC 1500  
Db 1537 GGCATGTGACAGCGACACCAAGCCTTGGCGGAGCTTCTTTGTGTCCGTGATCAAAAC 1596  
Qy 1501 AGCGACAGCAAAATTTATTTGAAATTCGATGAAAACAAGAAAGCAACCTTGGGTGAAA 1560  
Db 1597 AGCGACAGCAAAATTTATTTGAAATTCGATGAAAACAAGAAAGCAACCTTGGGTGAAA 1656  
Qy 1561 ACAAAGCTGAAAGCACTACTTTTAAACAACGCGCTGATGATTAAGTTATATCAAGGG 1620  
Db 1657 ACAAAGCTGAAAGCACTACTTTTAAACAACGCGCTGATGATTAAGTTATATCAAGGG 1716  
Qy 1621 CTTAATATCGTACTTATTTATTAAGAAATCTGAGCTCCGATGATTAAGTCTGTTA 1680  
Db 1717 CTTAATATCGTACTTATTTATTAAGAAATCTGAGCTCCGATGATTAAGTCTGTTA 1776  
Qy 1681 ACAAATCGAATTTGATTTGTGTCATATGAACATCATATGCGCAACAGAAACCTAGTT 1740  
Db 1777 ACAAATCGAATTTGATTTGTGTCATATGAACATCATATGCGCAACAGAAACCTAGTT 1836  
Qy 1741 TCACCAAGAAAGTACCAACCAACCAAGGTACTTACTTCAACAGGTGGCAAGGA 1800  
Db 1837 TCACCAAGAAAGTACCAACCAACCAAGGTACTTACTTCAACAGGTGGCAAGGA 1896  
Qy 1801 ATCTAGTTTACTTGAAGTGGCGAGCTTGTCTACTTATTTGAGAGTCTACTTGTCT 1860  
Db 1897 ATCTAGTTTACTTGAAGTGGCGAGCTTGTCTACTTATTTGAGAGTCTACTTGTCT 1956  
Qy 1861 AGACGTAGAAAGAAAGAAATGCT 1881  
Db 1957 AGACGTAGAAAGAAAGAAATGCT 1977

RESULT 2  
US-09-071-035-217  
Sequence 217, Application US/09071035  
Patent No. 6448043  
GENERAL INFORMATION:  
APPLICANT: Gil H. Choi  
TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides  
NUMBER OF SEQUENCES: 496  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Human Genome Sciences, Inc.  
STREET: 9410 Key West Avenue  
CITY: Rockville  
STATE: Maryland  
COUNTRY: USA  
ZIP: 20850  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage  
COMPUTER: HP Vectra 486/33  
OPERATING SYSTEM: MSDOS version 6.2  
SOFTWARE: ASCII Text  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/071,035  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
ATTORNEY/AGENT INFORMATION:  
NAME: A. Anders Brookes  
REGISTRATION NUMBER: 36,373  
REFERENCE/DOCKET NUMBER: PB369P2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (301) 309-8504  
TELEFAX: (301) 309-8512  
INFORMATION FOR SEQ ID NO: 217:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1983 base pairs  
TYPE: nucleic acid

STRANDEDNESS: double  
TOPOLOGY: linear  
US-09-071-035-217

Query Match 100.0%; Score 1881; DB 3; Length 1983;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1881; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 AAGAAGCAATTAAGGATTTGATACACCGTTAGTACCTTGTACTAATTTTGGCACTT 60
DB 100 AAGAAGCAATTAAGGATTTGATACACCGTTAGTACCTTGTACTAATTTTGGCACTT 159
QY 61 TTCACAGTGTATTAAGGACCAACAATGATTTGACAGAGAAAATGGGAGAGCGACAG 120
DB 160 TTCACAGTGTATTAAGGACCAACAATGATTTGACAGAGAAAATGGGAGAGCGACAG 219
QY 121 CTCGTGATTCACAAAAGAAAATGACGATTTACAGATCCGCTTAATTCAAAATAGCGGG 180
DB 220 CTCGTGATTCACAAAAGAAAATGACGATTTACAGATCCGCTTAATTCAAAATAGCGGG 279
QY 181 AAGAAGAGGAGGATTTGATTAATATCAAGATCTGGCAGATGATGACGTTTATTTAT 240
DB 280 AAGAAGAGGAGGATTTGATTAATATCAAGATCTGGCAGATGATGACGTTTATTTAT 339
QY 241 AACGTGACGAACGAATTTTACAGACCAAGGCGGACAGCGCAAGCGTTGATGACGATAA 300
DB 340 AACGTGACGAACGAATTTTACAGACCAAGGCGGACAGCGGCAAGCGTTGATGACGATAA 399
QY 301 CAAGCTGTCCAAAGTTTAACTCTGCGGAAACCTGTTGCTCAAGAAACACCGATGCAAT 360
DB 400 CAAGCTGTCCAAAGTTTAACTCTGCGGAAACCTGTTGCTCAAGAAACACCGATGCAAT 459
QY 361 GGAATGTGCACTGTTCAGTTACTTAATAAAACAAAATGTGAAGATGACGTATACAT 420
DB 460 GGAATGTGCACTGTTCAGTTACTTAATAAAACAAAATGTGAAGATGACGTATACAT 519
QY 421 AAGAAGAACCAAAAGAGGTGTAGTGTCTGCTACGAATATGTTGGTGGCTTCCAGTT 480
DB 520 AAGAAGAACCAAAAGAGGTGTAGTGTCTGCTACGAATATGTTGGTGGCTTCCAGTT 579
QY 481 TACGAATGATCAGCAACAGATGTTCTTAATAATATGAAACAGAAATTAACGGTT 540
DB 580 TACGAATGATCAGCAACAGATGTTCTTAATAATATGAAACAGAAATTAACGGTT 639
QY 541 GTTCAATTTTATCCCTTAATAATGTGTAGCAATGATGATTTTCAATGTGAAAAAGTA 600
DB 640 GTTCAATTTTATCCCTTAATAATGTGTAGCAATGATGATTTTCAATGTGAAAAAGTA 699
QY 601 GGAATGCTGAATAATGAAGATTTAATAGCGCAGAAATTTGTAATTTCTAAAAGCGAAGC 660
DB 700 GGAATGCTGAATAATGAAGATTTAATAGCGCAGAAATTTGTAATTTCTAAAAGCGAAGC 759
QY 661 TCACCAAGCAGATTAATAATATATCCAGAGATCAAAAGATGATTAATATCATGACAAAG 720
DB 760 TCACCAAGCAGATTAATAATATATCCAGAGATCAAAAGATGATTAATATCATGACAAAG 819
QY 721 GATTAAGAAACAAGCAAAACGCTTTATTAATGAGAAAAGTATGAATTTGGCAAAATGAT 780
DB 820 GATTAAGAAACAAGCAAAACGCTTTATTAATGAGAAAAGTATGAATTTGGCAAAATGAT 879
QY 781 TTCACAGAGAGAGATGAAACGAGAGAAATTAACATGTTAAATCTTGAGGTTGTTG 840
DB 880 TTCACAGAGAGAGATGAAACGAGAGAAATTAACATGTTAAATCTTGAGGTTGTTG 939
QY 841 TATATTTTGAAGAGTAAGGATCAATAATATGCAATTAATTAATGAATCAACAAA 900
DB 940 TATATTTTGAAGAGTAAGGATCAATAATATGCAATTAATTAATGAATCAACAAA 999
QY 901 ACACCATTTACAAATGAAGCAAAACAATCAAAACCTGTTGAAAACAGTCAAAATAT 960
DB 1000 ACACCATTTACAAATGAAGCAAAACAATCAAAACCTGTTGAAAACAGTCAAAATAT 1059
QY 961 ACCTCTAAAGTTGATTAACAAACCAAGCTTATGATGTTAAGATGTGGCAATTTGGC 1020
```

```
DB 1060 ACCTCTAAAGTTGATTAACAAACCAAGCTTATGATGTTAAGATGTGGCAATTTGGC 1119
QY 1021 AAAATTAATATCAAAATTTCTGTAATATTTCCATTTGGGATTTGACAGCAAAAGAGCGAC 1080
DB 1120 AAAATTAATATCAAAATTTCTGTAATATTTCCATTTGGGATTTGACAGCAAAAGAGCGAC 1179
QY 1081 GCTAATTAATATCAAAATTTCTGTAATATTTCCATTTGGGATTTGACAGCAAAAGAGCGAC 1140
DB 1180 GCTAATTAATATCAAAATTTCTGTAATATTTCCATTTGGGATTTGACAGCAAAAGAGCGAC 1239
QY 1141 AACGTGATCTTGGAGAGATATGCTTATATGATGATGAGGATATGATGATGATGATGAT 1200
DB 1240 AACGTGATCTTGGAGAGATATGCTTATATGATGATGAGGATATGATGATGATGATGAT 1299
QY 1201 GAAAATTAATCAAGTATGAAACAGCAAAATGCTTCACTGTCGCGTTAATCAAGCTAT 1260
DB 1300 GAAAATTAATCAAGTATGAAACAGCAAAATGCTTCACTGTCGCGTTAATCAAGCTAT 1359
QY 1261 ATTCTAGGCTTAACACAGGCGGCAACCTAAATTCGTTACTTTATGATTAATGA 1320
DB 1360 ATTCTAGGCTTAACACAGGCGGCAACCTAAATTCGTTACTTTATGATTAATGA 1419
QY 1321 AAGAAGATCTTACGAAGGCTTTAAATAAGAGGAAATGTTGATTAACGATACCGAC 1380
DB 1420 AAGAAGATCTTACGAAGGCTTTAAATAAGAGGAAATGTTGATTAACGATACCGAC 1479
QY 1381 GACCAACACCAACCACTGTTGAAGTTGACAGGTGGGAAACGTTTCAATTAAGTCAAT 1440
DB 1480 GACCAACACCAACCACTGTTGAAGTTGACAGGTGGGAAACGTTTCAATTAAGTCAAT 1539
QY 1441 GCGGATGTGACAGGACCAAGCCTTGGCGGAGCTTCCTTGTGTCGCGATCAAAAC 1500
DB 1540 GCGGATGTGACAGGACCAAGCCTTGGCGGAGCTTCCTTGTGTCGCGATCAAAAC 1599
QY 1501 AGCGACAGCAAAATTAATTTGAAATTCATGAAACAAAGCAAGCAACTTGGTGAAA 1560
DB 1600 AGCGACAGCAAAATTAATTTGAAATTCATGAAACAAAGCAAGCAACTTGGTGAAA 1659
QY 1561 ACAAAAGCTGAAGCACTTCTTTTCAACACAGCTGATGATTAAGTTATCAACAGG 1620
DB 1660 ACAAAAGCTGAAGCACTTCTTTTCAACACAGCTGATGATTAAGTTATCAACAGG 1719
QY 1621 CTTAATAGGATCTAATTAATTTTGAAGAACTGATGCTCCGATGATTAATGCTTGT 1680
DB 1720 CTTAATAGGATCTAATTAATTTTGAAGAACTGATGCTCCGATGATTAATGCTTGT 1779
QY 1681 ACAATTCGATTTGATTTGTGCTCAATGAAACATATGCAACAGAAACCTAGTT 1740
DB 1780 ACAATTCGATTTGATTTGTGCTCAATGAAACATATGCAACAGAAACCTAGTT 1839
QY 1741 TCACCAAAAAAGTACCAACAAACAAAGGATCTTACTTCAACAGGTGGCAAGGA 1800
DB 1840 TCACCAAAAAAGTACCAACAAACAAAGGATCTTACTTCAACAGGTGGCAAGGA 1899
QY 1801 ATCTAGCTTCTTAAGAAAGTGGCGAGCTTGTCACTTAATGCGAGGATCTAATTTGCT 1860
DB 1900 ATCTAGCTTCTTAAGAAAGTGGCGAGCTTGTCACTTAATGCGAGGATCTAATTTGCT 1959
QY 1861 AGACGTAGAAAAAATATGCT 1881
DB 1960 AGACGTAGAAAAAATGCT 1980
```

RESULT 3  
US-10-206-576-217  
; Sequence 217, Application us/10206576  
; Patent No. 6913907  
GENERAL INFORMATION:  
APPLICANT: Choi et al.  
TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides  
NUMBER OF SEQUENCES: 497  
CORRESPONDENCE ADDRESS:

ADDRESS: Human Genome Sciences, Inc.  
STREET: 9410 Key West Avenue  
CITY: Rockville  
STATE: Maryland  
COUNTRY: USA  
ZIP: 20850  
COMPUTER READABLE FORM:  
MEDIUM TYPE: CD-R  
COMPUTER: Dell Latitude  
OPERATING SYSTEM: Windows 98  
SOFTWARE: ASCII Text  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/206,576  
FILING DATE: 29-Jul-2002  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 09/071,035  
FILING DATE: 1998-05-04  
APPLICATION NUMBER: US 60/046,655  
FILING DATE: 1997-05-16  
APPLICATION NUMBER: US 60/044,031  
FILING DATE: 1997-05-06  
APPLICATION NUMBER: US 60/066,009  
FILING DATE: 1997-11-14  
ATTORNEY/AGENT INFORMATION:  
NAME: Hyman, Mark J.  
REGISTRATION NUMBER: 46,789  
REFERENCE/DOCKET NUMBER: PB369PDI1  
INFORMATION FOR SEQ ID NO: 217:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1983 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
SEQUENCE DESCRIPTION: SEQ ID NO: 217:  
US-10-206-576-217  
Query Match 100.0%; Score 1881; DB 3; Length 1983;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1881; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAAGCAATTAATAAAGTTGGTACACCGTTAGTACCTTGTACTTAATTTGGCCACTT 60  
DB 100 ATGAAGCAATTAATAAAGTTGGTACACCGTTAGTACCTTGTACTTAATTTGGCCACTT 159  
QY 61 TTCAACAAGTGTATTTAGGGACAACAATGCAATTTGGCAAAATTTGGGAGAGCGCAG 120  
DB 160 TTCAACAAGTGTATTTAGGGACAACAATGCAATTTGGCAAAATTTGGGAGAGCGCAG 219  
QY 121 CTCGTGATTCACAAAAGAAATGACGATTTACAGATCCGCTTATTCAAAATAGCGGG 180  
DB 220 CTCGTGATTCACAAAAGAAATGACGATTTACAGATCCGCTTATTCAAAATAGCGGG 279  
QY 181 AAAAAGATGAGCGAGTTTGAATAATATCAAGACCTGGCAGATGTGACGTTTAGTATTAT 240  
DB 280 AAAAAGATGAGCGAGTTTGAATAATATCAAGACCTGGCAGATGTGACGTTTAGTATTAT 339  
QY 241 AACGTGACGAAGATTTTACGAGCAAGAGCGGCGCAAGCCTTGAATGCACTATAA 300  
DB 340 AACGTGACGAAGATTTTACGAGCAAGAGCGGCGCAAGCCTTGAATGCACTATAA 399  
QY 301 CAAGCTGTCAGAAAGTTTAACTCTGGGAAAACCTGTTGCTCAAGGAACACCGATGCAAT 360  
DB 400 CAAGCTGTCAGAAAGTTTAACTCTGGGAAAACCTGTTGCTCAAGGAACACCGATGCAAT 459  
QY 361 GGGAAATGCTACTGTTCACTTAATAAATAAATAATGTTAAAGATGCACTATACATT 420  
DB 460 GGGAAATGCTACTGTTCACTTAATAAATAAATAATGTTAAAGATGCACTATACATT 519  
QY 421 AAAAGAAACCAAAAGAGGTGATGTTGCTGCTACGAATATGTTGCTGCTCCAGTT 480  
DB 520 AAAAGAAACCAAAAGAGGTGATGTTGCTGCTACGAATATGTTGCTGCTCCAGTT 579

QY 481 TACGAATGATCAAGCAAAAGATGTTCTCTATAAATATGGAACAGAAATTAACGGTT 540  
DB 580 TACGAATGATCAAGCAAAAGATGTTCTCTATAAATATGGAACAGAAATTAACGGTT 639  
QY 541 GTTCATATTTATCTTAAATAATGTGTAGCCAAATGATGTTTACATGTGAAAAAGTA 600  
DB 640 GTTCATATTTATCTTAAATAATGTGTAGCCAAATGATGTTTACATGTGAAAAAGTA 699  
QY 601 GGAATGCTGAAAAATGAAAGATTAATATGCGGAGAAATTTGTTATTTCTAAAAAGC 660  
DB 700 GGAATGCTGAAAAATGAAAGATTAATATGCGGAGAAATTTGTTATTTCTAAAAAGC 759  
QY 661 TCACGAGCAAGTAAATATATCCAGAGAGTCAAGAATGATTTATATCATGACGCAACG 720  
DB 760 TCACGAGCAAGTAAATATATCCAGAGAGTCAAGAATGATTTATATCATGACGCAACG 819  
QY 721 GATTAAGAACAGCAAAACGCTTTATTAATGCGGAAAAGTTATGAAATTTGGGAAAATGAT 780  
DB 820 GATTAAGAACAGCAAAACGCTTTATTAATGCGGAAAAGTTATGAAATTTGGGAAAATGAT 879  
QY 781 TTCAAGAGAGCAAGAAATGGAACGGGAAATTAACAGTTAAAAATCTTGAAGTTGCTCG 840  
DB 880 TTCAAGAGAGCAAGAAATGGAACGGGAAATTAACAGTTAAAAATCTTGAAGTTGCTCG 939  
QY 841 TATATTTTGAAGAAAGTAAAGCTCCAAATATAGCAGATTAATTAATGAATCAACAAA 900  
DB 940 TATATTTTGAAGAAAGTAAAGCTCCAAATATAGCAGATTAATTAATGAATCAACAAA 999  
QY 901 ACACCAATTTACATTTGAAGCAAAATCAATCAACACTGTTGAAAAACAGTCAAAATGAT 960  
DB 1000 ACACCAATTTACATTTGAAGCAAAATCAACACTGTTGAAAAACAGTCAAAATGAT 1059  
QY 961 ACCCTAAAGTTGATTAACAAACCAACAGCTTAGATGTTAAAGATGAGCAATTTGGCAG 1020  
DB 1060 ACCCTAAAGTTGATTAACAAACCAACAGCTTAGATGTTAAAGATGAGCAATTTGGCAG 1119  
QY 1021 AAAATTAATATCAAAATTTCTGTAAATATTCATTTGGGATTTGACAAAGAGCGAC 1080  
DB 1120 AAAATTAATATCAAAATTTCTGTAAATATTCATTTGGGATTTGACAAAGAGCGAC 1179  
QY 1081 GCTAATTAATACGTCAATTTCAATTTAGTTGATTAACATGATGACGCTTAACTTTGAT 1140  
DB 1180 GCTAATTAATACGTCAATTTCAATTTAGTTGATTAACATGATGACGCTTAACTTTGAT 1239  
QY 1141 AACGTGACTTCGAGAGATGATCTTATGCGTTATATGATGAGGATACAGTATGCTCCT 1200  
DB 1240 AACGTGACTTCGAGAGATGATCTTATGCGTTATATGATGAGGATACAGTATGCTCCT 1299  
QY 1201 GAAAAATTAATCAAGTGAATGAAACAAAGCAATGGCTTCACTGCGCGTTAATCCAGCTAT 1260  
DB 1300 GAAAAATTAATCAAGTGAATGAAACAAAGCAATGGCTTCACTGCGCGTTAATCCAGCTAT 1359  
QY 1261 ATTTCTAGCTTAACACAGGCGGCACTAAATTCGTTTACTTTATATGATTTAAATGAA 1320  
DB 1360 ATTTCTAGCTTAACACAGGCGGCACTAAATTCGTTTACTTTATATGATTTAAATGAA 1419  
QY 1321 AAAAGCAATCTTACGAAAGGCTTTTAAAAATGAGCGAAATGTTGATTAACGCTATACCGAC 1380  
DB 1420 AAAAGCAATCTTACGAAAGGCTTTTAAAAATGAGCGAAATGTTGATTAACGCTATACCGAC 1479  
QY 1381 GACCAAAACCAACCAATGTTGAAGTTGACAGGTGGGAAAAGTTTCAATTAAGTCGAT 1440  
DB 1480 GACCAAAACCAACCAATGTTGAAGTTGACAGGTGGGAAAAGTTTCAATTAAGTCGAT 1539  
QY 1441 GGGATGTGACAGCGACACACCTTTGGGGAGCTTCTTGTGTCCTGATCAAAAC 1500  
DB 1540 GGGATGTGACAGCGACACACCTTTGGGGAGCTTCTTGTGTCCTGATCAAAAC 1599  
QY 1501 AGCGACACAGCAAAATTTATTTGAAATTCGATGAAACAAAGAGCAACCTTGGGTGAAA 1560  
DB 1600 AGCGACACAGCAAAATTTATTTGAAATTCGATGAAACAAAGAGCAACCTTGGGTGAAA 1659  
QY 1561 ACAAAAGCTGAAGCAACTTATTTTAAACAAACGCGCTGATGATTAAGTTGATATCACAGGG 1620

| Accession | Sequence                                                          | Length |
|-----------|-------------------------------------------------------------------|--------|
| Db        | 1660 ACAAAGGTGAGCAACTCTTTTACAAACACGGCTGATGATTAAGTTGATATCACAGGG    | 1719   |
| Qy        | 1621 CTTAAATTCGCTACCTATTATTTAGAAAGAACTGAGCTCCGATGATTAATGCTCTTGTA  | 1680   |
| Db        | 1720 CTTAAATTCGCTACCTATTATTTAGAAAGAACTGAGCTCCGATGATTAATGCTCTTGTA  | 1779   |
| Qy        | 1681 ACAATTCGATTTGAATTTTGCTCATGATACATCATATGSCACAACGAAATCTAGTT     | 1740   |
| Db        | 1780 ACAATTCGATTTGAATTTTGCTCATGATACATCATATGSCACAACGAAATCTAGTT     | 1839   |
| Qy        | 1741 TCACAGAAAAAGTACCAACAAACACAAAGTACCTTCAACAGGTGGCAAAAGGA        | 1800   |
| Db        | 1840 TCACAGAAAAAGTACCAACAAACACAAAGTACCTTCAACAGGTGGCAAAAGGA        | 1899   |
| Qy        | 1801 ATCTACGTTTACTTAGAAGTGGCGACGCTCTTGCTACTTATTTGACGAGTCTACTTTGCT | 1860   |
| Db        | 1900 ATCTACGTTTACTTAGAAGTGGCGACGCTCTTGCTACTTATTTGACGAGTCTACTTTGCT | 1959   |
| Qy        | 1861 AGAGGTGAAAAGAAATGCT                                          | 1881   |
| Db        | 1960 AGAGGTGAAAAGAAATGCT                                          | 1980   |

## RESULT 4

```

US-09-071-035-219
; Sequence 219, Application US/09071035
; Patent No. 6448043
; GENERAL INFORMATION:
; APPLICANT: Gil H. Choi
; TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides
; NUMBER OF SEQUENCES: 496
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4MB storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/071,035
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: A. Anders Brooks
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB369P2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 219:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1687 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
US-09-071-035-219

```

|                       |                 |               |           |              |
|-----------------------|-----------------|---------------|-----------|--------------|
| Query Match           | 89.7%;          | Score 1687;   | DB 3;     | Length 1687; |
| Best Local Similarity | 100.0%;         | Pred. No. 0;  |           |              |
| Matches 1687;         | Conservative 0; | Mismatches 0; | Indels 0; | Gaps 0       |

Oy 96 AGAAGAAATGCGGAGAGCGACAGCTCGTGAATTCACAAAAAGAAATGACCGATTTTACC 155  
 |||||  
 Db 1 AGAAGAAATGCGGAGAGCGACAGCTCGTGAATTCACAAAAAGAAATGACCGATTTTACC 60

|    |      |                                                                 |      |
|----|------|-----------------------------------------------------------------|------|
| QY | 156  | IGATCCGCTTATTCATAAAATACCGGAAAAGAAATGAGCGAGTTGTATTAATATCAAGCAT   | 215  |
| Db | 61   | AGATCCGCTTATTCATAAAATACCGGAAAAGAAATGAGCGAGTTGTATTAATATCAAGCAT   | 120  |
| QY | 216  | GGCAGATGTGACGTTTAGTATTTAACTGTCACCAACGAATTTTACGACACGACGCG        | 275  |
| Db | 121  | GGCAGATGTGACGTTTAGTATTTAACTGTCACCAACGAATTTTACGACACGACGCG        | 180  |
| QY | 276  | AGGCGCAAGCGTTGATGCGCTAAACAAGCTGTCCAAAGTTTAACTCTGGGAAACCTGT      | 335  |
| Db | 181  | AGGCGCAAGCGTTGATGCGCTAAACAAGCTGTCCAAAGTTTAACTCTGGGAAACCTGT      | 240  |
| QY | 336  | TGCTCAAGGAACCAACCGATGCAAAATGGGAAATGTCACTGTTCAAGTACTTAATAAACAANA | 395  |
| Db | 241  | TGCTCAAGGAACCAACCGATGCAAAATGGGAAATGTCACTGTTCAAGTACTTAATAAACAANA | 300  |
| QY | 396  | TGCTAAAGATGCAAGTGTATATACATTAAAGAAAGAACCAAAAGAGGTGTAGTTGCTGTAC   | 455  |
| Db | 301  | TGCTAAAGATGCAAGTGTATATACATTAAAGAAAGAACCAAAAGAGGTGTAGTTGCTGTAC   | 360  |
| QY | 456  | GAATATGTGTGTGCGCTTCCAGATTACGAATATGATCAAGCAACAGATGTTTCTTATTA     | 515  |
| Db | 361  | GAATATGTGTGTGCGCTTCCAGATTACGAATATGATCAAGCAACAGATGTTTCTTATTA     | 420  |
| QY | 516  | ATTATGAAACAAGAAATTAGCGGTGTTCATATTTATCTTAAAAATGTGTAGCAATATGA     | 575  |
| Db | 421  | ATTATGAAACAAGAAATTAGCGGTGTTCATATTTATCTTAAAAATGTGTAGCAATATGA     | 480  |
| QY | 576  | TGCTATGTTTACATGTGAAAAAATGATGGAACCTGTGAAAATGAGAGATTAAATGCGCAGA   | 635  |
| Db | 481  | TGCTATGTTTACATGTGAAAAAATGATGGAACCTGTGAAAATGAGAGATTAAATGCGCAGA   | 540  |
| QY | 636  | ATTGTGTTATTTCTPAAAGCGAAGGCTCACCGGCAACGTAAATATATCCAAAGAGTCAA     | 695  |
| Db | 541  | ATTGTGTTATTTCTPAAAGCGAAGGCTCACCGGCAACGTAAATATATCCAAAGAGTCAA     | 600  |
| QY | 696  | AGATGATTTATATACATGACCAACGATTAAGAACCAAGCAAAACGCTTATTTACTGGGAA    | 755  |
| Db | 601  | AGATGATTTATATACATGACCAACGATTAAGAACCAAGCAAAACGCTTATTTACTGGGAA    | 660  |
| QY | 756  | AAGTTATGAAATTTGGCGAAAATGATTTCAACGAAGCAGAAATGGAACCGGAGAAATTAAC   | 815  |
| Db | 661  | AAGTTATGAAATTTGGCGAAAATGATTTCAACGAAGCAGAAATGGAACCGGAGAAATTAAC   | 720  |
| QY | 816  | AGTTAAAAAATTTGAGGTGTGTTGTTATTTTGAAGAAGTAAAGCTCCAAATTAATGC       | 875  |
| Db | 721  | AGTTAAAAAATTTGAGGTGTGTTGTTATTTTGAAGAAGTAAAGCTCCAAATTAATGC       | 780  |
| QY | 876  | AGAAATTATTTGAAATCAAAACAAACCAACATTTACATTTGAGAGCAAAACATCAAAACCC   | 935  |
| Db | 781  | AGAAATTATTTGAAATCAAAACAAACCAACATTTACATTTGAGAGCAAAACATCAAAACCC   | 840  |
| QY | 936  | TGTTGAAAAAAACATTCAAAAATGATTAATCTCTPAAAGTTGATTAACCAACCAACGCTTGA  | 995  |
| Db | 841  | TGTTGAAAAAAACATTCAAAAATGATTAATCTCTPAAAGTTGATTAACCAACCAACGCTTGA  | 900  |
| QY | 996  | TGCTAAAGATGTGCAATTGGCGAAAAATTTAAATATCAAAATTTCTGTAAATATTTCCATT   | 1055 |
| Db | 901  | TGCTAAAGATGTGCAATTGGCGAAAAATTTAAATATCAAAATTTCTGTAAATATTTCCATT   | 960  |
| QY | 1056 | GGGATTTGCAAGCAAAAGAGCGACGCTAATTAATACGTCAAAATTCATTTTAGTTGATTA    | 1115 |
| Db | 961  | GGGATTTGCAAGCAAAAGAGCGACGCTAATTAATACGTCAAAATTCATTTTAGTTGATTA    | 1020 |
| QY | 1116 | ACATGATGCAAGCTTAACTTTTGAATACGTGACTTGTGAGAGATGATCTTATGCGTTATA    | 1175 |
| Db | 1021 | ACATGATGCAAGCTTAACTTTTGAATACGTGACTTGTGAGAGATGATCTTATGCGTTATA    | 1080 |
| QY | 1176 | TGATGGGATATCAAGTATGCTCCTGAAAATTTATCAATGATCTGAACAAGCAAAATGCGTT   | 1235 |
| Db | 1081 | TGATGGGATATCAAGTATGCTCCTGAAAATTTATCAATGATCTGAACAAGCAAAATGCGTT   | 1140 |



```

QY 936 TGTGAAAAACGTCATAAATGATACCTTAAGTGAATTAACAACCAAGCTTAG 995
DB 841 TGTGAAAAACGTCATAAATGATACCTTAAGTGAATTAACAACCAAGCTTAG 900
QY 996 TGTGAAAAACGTCATAAATGATACCTTAAGTGAATTAACAACCAAGCTTAG 1055
DB 901 TGTGAAAAACGTCATAAATGATACCTTAAGTGAATTAACAACCAAGCTTAG 960
QY 1056 GGGGATTCGACAGCAAGGAGCGGCTTAATTAATACGTCAAATTCATTTAGTATTA 1115
DB 961 GGGGATTCGACAGCAAGGAGCGGCTTAATTAATACGTCAAATTCATTTAGTATTA 1020
QY 1116 ACATGATGAGCGCTTAATTCATTTAGTATTAATGATACCTTAAGTGAATTAACAACCAAGCTTAG 1175
DB 1021 ACATGATGAGCGCTTAATTCATTTAGTATTAATGATACCTTAAGTGAATTAACAACCAAGCTTAG 1080
QY 1176 TGTGAAAAACGTCATAAATGATACCTTAAGTGAATTAACAACCAAGCTTAG 1235
DB 1081 TGTGAAAAACGTCATAAATGATACCTTAAGTGAATTAACAACCAAGCTTAG 1140
QY 1236 CACTGTCGCGCTTAATTCAGCGGTAATTCCTTACGCTTAACAACGAGCGGCACTTAAATTT 1295
DB 1141 CACTGTCGCGCTTAATTCAGCGGTAATTCCTTACGCTTAACAACGAGCGGCACTTAAATTT 1200
QY 1296 CGTTTACCTTAATTCAGCTTAATTAATGATACCTTAAGTGAATTAACAACCAAGCTTAG 1355
DB 1201 CGTTTACCTTAATTCAGCTTAATTAATGATACCTTAAGTGAATTAACAACCAAGCTTAG 1260
QY 1356 GAATGTTGTAACGTCATACGCAAGCAACCAACCACTGTTGAAGTTGTGACAG 1415
DB 1261 GAATGTTGTAACGTCATACGCAAGCAACCAACCACTGTTGAAGTTGTGACAG 1320
QY 1416 TGGGAAACGTTTCAATTAATGATGAGCGATGTGACAGCGGCAACAGCTTGGCGGAGC 1475
DB 1321 TGGGAAACGTTTCAATTAATGATGAGCGATGTGACAGCGGCAACAGCTTGGCGGAGC 1380
QY 1476 TTTCTTTGTCGTCGCGATCAAAACGAGCAACAGAAATTTGTAATGATGAAAC 1535
DB 1381 TTTCTTTGTCGTCGCGATCAAAACGAGCAACAGAAATTTGTAATGATGAAAC 1440
QY 1536 AACGAAGCAGCACTTGGGTGTAACAAAGCTGAAGCACTTCTTACCAACGAGC 1595
DB 1441 AACGAAGCAGCACTTGGGTGTAACAAAGCTGAAGCACTTCTTACCAACGAGC 1500
QY 1596 TGTGATTTAGTTGATATCAACGAGCTTAATTCGCTTATTTAGAAAGAACTGT 1655
DB 1501 TGTGATTTAGTTGATATCAACGAGCTTAATTCGCTTATTTAGAAAGAACTGT 1560
QY 1656 AGCTCTGATGATTAATGCTTGTGTAACAAATGGGATTTGTCATGAAACAATC 1715
DB 1561 AGCTCTGATGATTAATGCTTGTGTAACAAATGGGATTTGTCATGAAACAATC 1620
QY 1716 ATATGCAACAAGAAACCTTATTCACGAGAAAGTACCAACCAACCAAGGATC 1775
DB 1621 ATATGCAACAAGAAACCTTATTCACGAGAAAGTACCAACCAACCAAGGATC 1680
QY 1776 CTTACCT 1782
DB 1681 CTTACCT 1687

```

```

RESULT 6
US-09-107-532A-3598
; Sequence 3598, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESS: GENOME THERAPEUTICS CORPORATION

```

```

? STREET: 100 Beaver Street
? CITY: Waltham
? STATE: Massachusetts
? COUNTRY: USA
? ZIP: 02354
? COMPUTER READABLE FORM:
? MEDIUM TYPE: CD-ROM ISO9660
? OPERATING SYSTEM: <Unknown>
? SOFTWARE: ASCII
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/09/107,532A
? FILING DATE: 30-Jun-1998
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: 60/085,598
? FILING DATE: 14 May 1998
? APPLICATION NUMBER: 60/051,571
? FILING DATE: July 2, 1997
? ATTORNEY/AGENT INFORMATION:
? NAME: Ariniello, Pamela Deneke
? REGISTRATION NUMBER: 40,489
? REFERENCE/DOCKET NUMBER: GTC-012
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (781)893-5007
? TELEFAX: (781)893-8277
? INFORMATION FOR SEQ ID NO: 3598:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 1995 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: double
? TOPOLOGY: circular
? MOLECULE TYPE: DNA (genomic)
? HYPOTHETICAL: NO
? ANTI-SENSE: NO
? ORIGINAL SOURCE:
? ORGANISM: Enterococcus faecium
? FEATURE:
? NAME/KEY: misc_feature
? LOCATION: (8) LOCATION 1..1995
? SEQUENCE DESCRIPTION: SEQ ID NO: 3598:
US-09-107-532A-3598

Query Match      53.9%; Score 1014.6; DB 3; Length 1995;
Best Local Similarity 71.9%; Pred. No. 5.9e-238;
Matches 1342; Conservative 0; Mismatches 519; Indels 6; Gaps 1;

QY 1 ATGAAGCAATTAATAAAGTTGGTACACCGTTAGTACCTTGTACTTAATTTGGCACTT 60
DB 118 ATGAATAATCATTAATAAATAACGTTAGTGAAGTCTTTCTTCTTATTTACATTA 177
QY 61 TTCAAGTGTATTAGGACAAACAACGTCATTTGCAGAAAGAAATGGGAGAGCGACAG 120
DB 178 CTCACAAACAGCTTGGCGGCAAAAAGTGTGACAGAGACAG-----CAGCTCAA 231
QY 121 CTCGATTCACAAAAGAAATGACGATTTACAGATCCGCTTATTCAAAATGCGGG 180
DB 232 GTCACTCTTCATTAAGAAATGACTGATTTACCGGATCTTTATTCAAAACGCGGG 291
QY 181 AAGAAGATGAGCGATTTGATTAATATCAAGGATCGGAGATGTGACGTTAGTATTAT 240
DB 292 AAGAAGATGAGCGATTTGATTAATATCAAGGATCGGAGATGTGACGTTAGTATTAT 351
QY 241 AACGTGACAGCAATTTTACAGACAGAGCGGCAAGCGTGAATGACGTTAA 300
DB 352 AACGTGACAGCAATTTTATGCGCAAGAGATTAAGAGCGTCCGTGATCAGCAAAA 411
QY 301 CAGCTGTCGCAAGTTTAACTCTGGGAAACCTGTTGCTCAAGAAACACCGATGCAAT 360
DB 412 CAGCAGTCGCACTTTGATCTCTGTACACAGATGTCCTTCAAGAAACAGATGCTGAT 471
QY 361 GGAATGTCACTGTTCACTTACCTTAATAAACAATGTGAAGATGAGTATACCAT 420
DB 472 GGAATGTCACTTATCTTACCTTAATAAACAATGTGAAGATGAGTATACGATC 531

```











Db 10614 ATGATGAGAAAAAGAAATTCGTGGAATGGAACCTGTATTAAACAGAAAGCCCTCTGCTC 10555  
Qy 725 AAGACAGACAAACCCCTTATTACTGGGAAAAGTTATGAAATTCGCGAAATGATTTCA 784  
Db 10554 GAGAACAAATTCACATATTATTCACAGATTAATAAAGCAATTTGAAGCGAATATTCAAG 10495  
Qy 785 CAGAACAGAGATGGAACGGGAGAAATTAAACATTAAAACTTGAGGTGGTTCGTATTA 844  
Db 10494 CAACGGTAGAAGAAAGAAATAGTATATTAGCAGAGTTACAAATTTTATGACACTGCTA 10435  
Qy 845 TTTTAGAAGAAATGAAAGCTCCAAATATATGAGAAATTAATTTGAAAATCAACAAACAC 904  
Db 10434 TTGACAAATTGATCAAGATCGTAGCAATGCAAGTTGATTAAGACATCTTTAAATC 10375  
Qy 905 CATTTCAAATTGAGCAACAAATCAACACCTGTTGAAAAAAGTCGCAAAAATGATACCT 964  
Db 10374 TACAAACAAATACATGATTGATGTACATCTTATTAATAAGCAGATGCTGAAAAACGA 10315  
Qy 965 CTAAAGTTGAT 975  
Db 10314 TTAATGATGAT 10304

## RESULT 12

US-08-781-986A-171/c  
Sequence 171, Application US/08781986A  
Patent No. 6737248  
GENERAL INFORMATION:  
APPLICANT: Charles Kuech  
TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences  
NUMBER OF SEQUENCES: 5255  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Human Genome Sciences, Inc.  
STREET: 9410 Key West Avenue  
CITY: Rockville  
STATE: Maryland  
COUNTRY: USA  
ZIP: 20850  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage  
COMPUTER: HP Vectra 486/33  
OPERATING SYSTEM: MSDOS version 6.2  
SOFTWARE: ASCII Text  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/781,986A  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Benson, Bob  
REGISTRATION NUMBER: 30,446  
REFERENCE/DOCKET NUMBER: PB248PP  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (301) 309-8504  
TELEFAX: (301) 309-8512  
INFORMATION FOR SEQ ID NO: 171:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 11126 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
US-08-781-986A-171

Query Match 3.0%; Score 56.6; DB 3; Length 11126;

Best Local Similarity 45.7%; Pred. No. 0.0011;

Matches 197; Conservative 0; Mismatches 234; Indels 0; Gaps 0;  
Qy 545 ATATTATCTTAAATGTGTGAGCCAAATGATGATGTTACATGTGAAAAAGTAGGAA 604  
Db 10734 ATAACGACGACAGATGAAGAAACAAATATTGCAATAGCACAGATTGAAAAAGATTAA 10675

Qy 605 CTGCTGAAATGAGGATTTAAATGGCGGAATTTGTATTCTTAATAAGGAGCTCAC 664  
Db 10674 TTAACCTTAAACAAATTTGGCTAGTGAAGTGAATGATGAGATGAGCAATATTATTCG 10615  
Qy 665 CAGGCACAGTAAATATATTCACAGAGAGTCAAAAGATGATTAATATACATGACAAACGATTA 724  
Db 10614 ATGATGAGAAAAAGAAATTTCCGTAAATGAAACCTTTTATTAAACAGAAAGCCCTCTGCTC 10555  
Qy 725 AAGAACAGCAAAACGCTTTATTACTGGGAAAAGTTATGAAATTCGCGAAATGATTTCA 784  
Db 10554 GAGAACAAATTCACATATTATTCACAGATTAATAAAGCAATTTGAAGCGAATATTCAAG 10495  
Qy 785 CAGAACAGAGATGGAACGGGAGAAATTAAACATTAAAACTTGAGGTGGTTCGTATTA 844  
Db 10494 CAACGGTAGAAGAAAGAAATAGTATATTAGCAGAGTTACAAATTTTATGACACTGCTA 10435  
Qy 845 TTTTAGAAGAAATGAAAGCTCCAAATATATGAGAAATTAATTTGAAAATCAACAAACAC 904  
Db 10434 TTGACAAATTGATCAAGATCGTAGCAATGCAAGTTGATTAAGACATCTTTAAATC 10375  
Qy 905 CATTTCAAATTGAGCAACAAATCAACACCTGTTGAAAAAAGTCGCAAAAATGATACCT 964  
Db 10374 TACAAACAAATACATGATTGATGTACATCTTATTAATAAGCAGATGCTGAAAAACGA 10315  
Qy 965 CTAAAGTTGAT 975  
Db 10314 TTAATGATGAT 10304

## RESULT 13

US-08-916-421B-1/c  
Sequence 1, Application US/08916421B  
Patent No. 6503729  
GENERAL INFORMATION:  
APPLICANT: Bull et al.  
TITLE OF INVENTION: Complete Genome Sequence of the Methanogenic Archaeon, Methanoco-  
Patent No. 6503729  
TITLE OF INVENTION: jamaaschi1  
FILE REFERENCE: PB275  
CURRENT APPLICATION NUMBER: US/08/916,421B  
CURRENT FILING DATE: 1997-08-22  
PRIOR APPLICATION NUMBER: US 60/024,428  
PRIOR FILING DATE: 1996-08-22  
NUMBER OF SEQ ID NOS: 3  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 1  
LENGTH: 1664976  
TYPE: DNA  
ORGANISM: Methanococcus jamaaschi1  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (28222)..(28222)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature  
LOCATION: (28257)..(28258)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature  
LOCATION: (84773)..(84773)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature  
LOCATION: (84808)..(84808)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature  
LOCATION: (84812)..(84812)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature  
LOCATION: (98120)..(98120)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature  
LOCATION: (98159)..(98159)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature

```
LOCATION: (98239) .. (98239)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (98266) .. (98266)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (98343) .. (98343)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (10398) .. (10398)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (148948) .. (148948)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (163385) .. (163385)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (19189) .. (19189)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (19195) .. (19195)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (23180) .. (23180)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (234187) .. (234187)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (234220) .. (234220)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (234614) .. (234614)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (309398) .. (309398)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (309418) .. (309418)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (312837) .. (312837)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (312993) .. (312993)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (319226) .. (319226)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (559167) .. (559167)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (559241) .. (559241)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (600992) .. (600992)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (622708) .. (622708)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (657081) .. (657081)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (657203) .. (657203)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (674435) .. (674435)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (682442) .. (682442)
```

```
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (713552) .. (713552)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (741684) .. (741684)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (779455) .. (779455)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (779676) .. (779676)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (85539) .. (85539)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (871619) .. (871619)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1084830) .. (1084830)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1096846) .. (1096846)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1119881) .. (1119881)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1130881) .. (1130881)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1310988) .. (1310988)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1313224) .. (1313224)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1349473) .. (1349473)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1470091) .. (1470091)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1569020) .. (1569020)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1602912) .. (1602912)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1603734) .. (1603734)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1637998) .. (1637998)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1664854) .. (1664854)
OTHER INFORMATION: n equals a, t, c, or g
US-08-916-421B-1
```

Query Match 3.0%; Score 56.4; DB 3; Length 1664976;  
Best Local Similarity 50.0%; Pred.No. 0.0071;  
Matches 141; Conservative 0; Mismatches 141; Indels 0; Gaps 0;

```
QY 833 TTGCTTCGATATTATTAGAGAGTAAAGCTCCAAATATGCAAGATTATGAAATC 892
Db 29123 TTGATTACAAACAAAGAGAGTAAATTAATAAAGAGTATATATG 29064
QY 893 AAACAAAACACCATTTACATTTGAAAGCAACATCAACACCTGTGAAAAACAGTCA 952
```

```

Db      29063 AAAAAAAAAATTAAATAAATACATAGATAGAAATATATNGATPAATAATTTGAGAAAGATGCTTTG 29004
QY      953  AAATGATACCTTAAAGTTGATATAAACACACCAAGCTTAGATGGTAAAGATGTGGCA 1012
Db      29003 TAAATCAAGCTATTAATATATATGAAAGAACCAAAATTTTTTTCAGCTTCAAGTGAAGCGAG 28944
QY      1013  TTGGCGAAAAAATTAAATATCAAAATTTCTGTAAATATTCATTGGGGATTTGCAGACAAG 1072
Db      28943 TTGGCTATATACCAAGCTATATATAAAATGAAGAACCAAAATCACTAAATTAATGTGAATT 28884
QY      1073  AAGCGACGCTAAATTAATACGTCAAAATTCAAATTTAGTTGATA 1114
Db      28883 ATGGCTACTCAAAATATATCTATATAACTTAATAGTTGTTAA 28842

RESULT 14
US-09-692-570-1/c
; Sequence 1, Application US/09692570
; Patent No. 6797466
; GENERAL INFORMATION:
; APPLICANT: Bult et al.
; TITLE OF INVENTION: Complete Genome Sequence of the Methanogenic Archaeon, Methanoco
; Patent No. 6797466
; TITLE OF INVENTION: jannaschii
; FILE REFERENCE: PB275C1
; CURRENT APPLICATION NUMBER: US/09/692,570
; CURRENT FILING DATE: 2003-01-14
; PRIOR APPLICATION NUMBER: US 60/024,428
; PRIOR FILING DATE: 1996-08-22
; PRIOR APPLICATION NUMBER: US 08/916,421
; PRIOR FILING DATE: 1997-08-22
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1664976
; TYPE: DNA
; ORGANISM: Methanococcus jannaschii
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (28222)..(28222)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (28257)..(28258)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (84773)..(84773)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (84808)..(84808)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (84812)..(84812)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (98120)..(98120)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (98159)..(98159)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (98239)..(98239)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (98266)..(98266)
; OTHER INFORMATION: n equals a, t, c, or g

```

|          |                        |                                |                                           |
|----------|------------------------|--------------------------------|-------------------------------------------|
| FEATURE: | NAME/KEY: misc_feature | LOCATION: (98343) .. (98343)   | OTHER_INFORMATION: n equals a, t, c, or g |
| FEATURE: | NAME/KEY: misc_feature | LOCATION: (103598) .. (103598) | OTHER_INFORMATION: n equals a, t, c, or g |
| FEATURE: | NAME/KEY: misc_feature | LOCATION: (146948) .. (146948) | OTHER_INFORMATION: n equals a, t, c, or g |
| FEATURE: | NAME/KEY: misc_feature | LOCATION: (163385) .. (163385) | OTHER_INFORMATION: n equals a, t, c, or g |
| FEATURE: | NAME/KEY: misc_feature | LOCATION: (191989) .. (191989) | OTHER_INFORMATION: n equals a, t, c, or g |
| FEATURE: | NAME/KEY: misc_feature | LOCATION: (191995) .. (191995) | OTHER_INFORMATION: n equals a, t, c, or g |
| FEATURE: | NAME/KEY: misc_feature | LOCATION: (231980) .. (231980) | OTHER_INFORMATION: n equals a, t, c, or g |
| FEATURE: | NAME/KEY: misc_feature | LOCATION: (234187) .. (234187) | OTHER_INFORMATION: n equals a, t, c, or g |
| FEATURE: | NAME/KEY: misc_feature | LOCATION: (234220) .. (234220) | OTHER_INFORMATION: n equals a, t, c, or g |
| FEATURE: | NAME/KEY: misc_feature | LOCATION: (234814) .. (234814) | OTHER_INFORMATION: n equals a, t, c, or g |
| FEATURE: | NAME/KEY: misc_feature | LOCATION: (309396) .. (309396) | OTHER_INFORMATION: n equals a, t, c, or g |
| FEATURE: | NAME/KEY: misc_feature | LOCATION: (309418) .. (309418) | OTHER_INFORMATION: n equals a, t, c, or g |
| FEATURE: | NAME/KEY: misc_feature | LOCATION: (312837) .. (312837) | OTHER_INFORMATION: n equals a, t, c, or g |
| FEATURE: | NAME/KEY: misc_feature | LOCATION: (312993) .. (312993) | OTHER_INFORMATION: n equals a, t, c, or g |
| FEATURE: | NAME/KEY: misc_feature | LOCATION: (319226) .. (319226) | OTHER_INFORMATION: n equals a, t, c, or g |
| FEATURE: | NAME/KEY: misc_feature | LOCATION: (559167) .. (559167) | OTHER_INFORMATION: n equals a, t, c, or g |
| FEATURE: | NAME/KEY: misc_feature | LOCATION: (559241) .. (559241) | OTHER_INFORMATION: n equals a, t, c, or g |
| FEATURE: | NAME/KEY: misc_feature | LOCATION: (600992) .. (600992) | OTHER_INFORMATION: n equals a, t, c, or g |

```
NAME/KEY: misc feature
LOCATION: (622708)..(622708)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (657081)..(657081)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (657203)..(657203)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (674435)..(674435)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (682442)..(682442)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (713652)..(713652)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (741684)..(741684)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (779455)..(779455)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (779676)..(779676)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (855539)..(855539)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (871619)..(871619)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (1084830)..(1084830)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (1096846)..(1096846)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (1119881)..(1119881)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (1130881)..(1130881)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (1310988)..(1310988)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (1313224)..(1313224)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (1349473)..(1349473)
OTHER INFORMATION: n equals a, t, c, or g
Query Match 3.0%; Score 56.4; DB 3; Length 1664976;
```

```
Best Local Similarity 50.0%; Pred. No. 0.0071;
Matches 141; Conservative 0; Mismatches 141; Indels 0; Gaps 0;

QY 833 TTGGTTCGTATATTTAGAGAGTAAGTAAGCTCCAAATTAAGCAGATTAATTTGAAATC 892
DB 29123 TTGATTACAAAACAAAGAGAGAGATGAATTAATTAAGAGATTATATG 29064
QY 893 AAACAAAAACACCATTTCAATTGAAGCAACATCAACACCTGTGAAAAACAGTCA 952
DB 29063 AAAAAATTTATATATCTATAGTAGAATATATGAATTTAGAGAGTGTCTTG 29004
QY 953 AAAATGATACCTCTAAAGTTGATTAACACCAAGCTTAGATGTAAGATGGCA 1012
DB 29003 TAAATGCAAGCTATTAATTAATGAAGAAAGAAATTTTACAGCTTCAAGTGAAGCAG 28944
QY 1013 TTGGCGAAAAATTAATATCAATTTCTGTAAATTTCCATGGGAGATGACACAAAG 1072
DB 28943 TTGGCTATATACCGCTATATTAATAAATGAACCAAGAACTAATAATATGGAATT 28884
QY 1073 AAGCGAGCTAATATATACGTCAATTCATTTAGTATG 1114
DB 28883 ATGGCTACTCAATATATCTATATACATTAGTCTGTATAAA 28842

RESULT 15
US-09-806-708B-22/c
; Sequence 22, Application US/09806708B
; Patent No. 6784342
; GENERAL INFORMATION:
; APPLICANT: The University of British Columbia
; TITLE OF INVENTION: Regulation of Embryonic Transcription in Plants
; FILE REFERENCE: 4810-58741
; CURRENT APPLICATION NUMBER: US/09/806,708B
; PRIOR FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: US 60/147,133
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 22
; LENGTH: 1141
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; NAME/KEY: promoter
; LOCATION: (1)..(1141)
; OTHER INFORMATION: consensus sequence of A.t., L.a., and B.n. FAEI promoters
US-09-806-708B-22

Query Match 2.9%; Score 54; DB 3; Length 1141;
Best Local Similarity 10.7%; Pred. No. 0.0022;
Matches 94; Conservative 330; Mismatches 449; Indels 9; Gaps 2;

QY 350 CCGATGCAATGGGATGCTACTGTTCAGTTACCTTAACAAAACAAATGTAAGATGACG 409
DB 944 MGNNNNNNNWTTDVRMAKAKKNNNNNNNAWTAQYRAATNNKMAATHMKWHGASHKRT 885
QY 410 TGTATACCATTAAGAAAGAGAGGAGGTGATGCTGCTACGATATATGATGCTG 469
DB 884 RHHTTCRTKTKNNNNNNNAKRYTYTHAARRMMAWTKTNNNNNNNNNNNNNNNNNNNNNN 825
QY 470 CGTTCACGATTACGAATGATCAAGCAACAGATGTTCTTAATATATGACACAGAAG 529
DB 824 BKHSHWNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 765
QY 530 AATTAGCGGTGTATATATTTATCTTAATAATGTGTAGCCAAATGATGATGTTAATCA 589
DB 764 THTDWCYTWMTWMTWMTWMTWMTWMTWMTWMTWMTWMTWMTWMTWMTWMTWMTWMTW 705
QY 590 TGAATAAGTGAAGTCTGTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 649
DB 704 AATATNNCGWNNNDPARKTNNNTTVVRRMMYTKRWSTTRHHYTGATNNNNNNNNNNNN 645
QY 650 AAAGGAAAGCTCACAG-----GCACAGTAAATATATCAAGAGTCAAGATGATGAT 703
```

```
Db 644 NNNNNNSCCTCTRMWMTMRWTKMGDGTVAKKVXWRDTCCTVYDVWADSWVWVYANMMCR 585
QY 704 TATATACATGACACACGATTAAGAACAAACGCTTATATCTGGAAGAAAGTTATG 763
Db 584 DVTYTRMNTYCKSYASHSYWYGSNNAMWYRYSARNWSMARWTTTR--NNMMWSGVRMR 528
QY 764 AAATTGGGGAATAATGTTTCACAGACGAGAGAAATGGAACGGAGAAATTAACAGTTAAA 823
Db 527 WAGTMMRRMMNNNTDTRYYMMKKRABRTTYYDSMCNAKSWRGNNRRAMMMWAAAN 468
QY 824 ATCTGAGGTGGTTCGTATATATTTAGAAAGTAAGCTCCAAATAATGACAAATTA 883
Db 467 DAGAMDHTWYMGNTMMWRBAMKMMWMAWCRBAYCCNNNNNRACVWHKHMMRWTKYM 408
QY 884 TTGAAAATCAAAACAAAACACATTTACATTTGAGCAAAACAATCAACAACCTGTGAAA 943
Db 407 WKAACNNNNBKAMMYRVAWMMYRSRDTTNTDMMWMTSDMBMHWYTVDTMMRAMNNNNNN 348
QY 944 AAACAGTCAAAAATGATACCTCTAAAGTGATTAACAACACCAAGCTTAGATGTAAG 1003
Db 347 WMBCKTTSWMMWMDHWNTHCTYGNNTWSGAYEBWAAWSMMAAGASNBVTYNNCMCRMTYMGK 288
QY 1004 ATGTGCAATTGGCGAAAAATTAATATCAATTTCTGTAATATTCATTGGGGATTG 1063
Db 287 TMTNNNNNNKAWYRTKTVAMCNRNRYVDTAWTBKRNKYCYAYBMYBYBMWKGHMBW 228
QY 1064 CAGACAAGAGGCGACGCTAATAATACGTCGAATTCAATTAGTTAGTTAACAATGATG 1123
Db 227 WRRABHRSWNNMMWVCYRNKYWVSWHYHMYRBYKMABAVGCNNNNWKDRMAHHHWCATNNNM 168
QY 1124 CAGCCTTAACCTTTGATTAACGAGCTTCGAGAGAGATGCTATGCTATATGATGGG 1183
Db 167 MMMWTAIYMHMHMKGAATNNKTABRDBHBAHVKTMYWRDYWCAMCMNNAKAKVRTA 108
QY 1184 ATACAGTATGCTCTGAAATAATTAACAAGTGAACAAG 1225
Db 107 MGHMYYTDRYVSANNTGVMMMMRWCMWYSMMNNRWYRMG 66
```

Search completed: July 5, 2006, 22:11:16  
Job time : 273 secs

**THIS PAGE BLANK (USPTO)**